

Education

CUK - H06051-28 - P8575

THE BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY

NOVEMBER 1989.

VOL. 59, PART 3



THE BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY

Edited by

Professor HAZEL FRANCIS

Assistant Editors:

Dr H. C. M. CARROLL

Prof. C. W. DESFORGES

Dr D. FONTANA

Dr F. J. TAYLOR

Sub-Editor:

Dr S. CUNNINGHAM

Book Review Editor:

Prof. C. W. DESFORGES

Advised by the Editorial Committee:

Professor G. BROWN

Ms. I. LUNT

Dr C. LEACH

Professor P. W. ROBINSON

Mr E. WILKINSON

Vol. 59, Part 3

The British Journal of Educational Psychology is issued by the British Psychological Society once a term, in February, June and November. The subscription is £25.00 a year (\$50.00 North America), or £8.50 each number.

Subscriptions and Orders should be sent to Scottish Academic Press, 139 Leith Walk, Edinburgh, EH6 8NS. Members of the British Psychological Society receive the *Journal* on special terms.

Papers for publication in duplicate should be sent to Professor Charles Desforges, University of Exeter, St Luke's, Heavitree Road, Exeter, EX1 2LU. Guidance to contributors on the preparation of papers for the *Journal* is given at the end of the November number each year.

All Books for Review for *The British Journal of Educational Psychology* should be sent to: Flora Macleod, University of Exeter, St Luke's, Heavitree Road, Exeter, EX1 2LU.

Advertisements for this *Journal* should be sent to Scottish Academic Press.

Published by

SCOTTISH ACADEMIC PRESS, 139 Leith Walk, Edinburgh, EH6 8NS,
for the British Journal of Educational Psychology, Ltd.

All rights reserved

© 1989. The British Journal of Educational Psychology

PRINTED IN GREAT BRITAIN BY
BELL AND BAIN LTD., GLASGOW

WHEN LESS INFORMATION IS MORE INFORMATIVE:
DIAGNOSING TEACHER EXPECTATIONS FROM BRIEF
SAMPLES OF BEHAVIOUR

By ELISHA BABAD
(Hebrew University of Jerusalem, Israel)
FRANK BERNIERI
(Oregon State University, USA)

AND
ROBERT ROSENTHAL
(Harvard University, USA)

SUMMARY. Teacher behaviour reflecting their differential expectations was investigated in a context-minimal method, where judges rate extremely brief (10-second) clips of videotaped teacher behaviour, separated into isolated non-verbal and verbal channels (face, body, speech content, tone of voice, etc.). Teachers were recorded when *talking about* and *talking to* high- and low-expectancy students. Contrary to recent claims that teacher expectancy effects are negligible and that teachers' differential behaviour is generally appropriate and reality-based, expectancy effects of substantial magnitude were found in this study, especially in affective and non-verbal behaviours.

Teachers were rated as showing more negative affect in the non-verbal channels, and as more dogmatic in the non-verbal and transcript channels, when talking about low expectancy compared to high expectancy students. When talking to students and teaching them briefly, facially communicated expectancy differences were found in ratings of negative affect and active teaching behaviour.

The findings supported a view of teachers as attempting to compensate low-expectancy students in controllable, direct teaching behaviours, at the same time transmitting (or "leaking") negative affect in less controllable, mostly non-verbal channels. It was also found that teachers who were more susceptible to biasing information were more negative and showed more intense expectancy effects than unbiased teachers in certain verbal and non-verbal channels.

INTRODUCTION

CAN more be learned about teachers by studying less of their behaviour? Current research in verbal and non-verbal communication indicates the utility and robustness of very brief samples of behaviour. Evidence from deception research (e.g., Zuckerman *et al.*, 1986), doctor-patient research (Milmoë *et al.*, 1967; Rosenthal *et al.*, 1984; Blanck, Rosenthal, Vannicelli, and Lee, 1986), and courtroom behaviour research (e.g., Blanck, 1987) shows the staggering amount of information in, and potential impact of, a few seconds of human behaviour. The intent of this study is to demonstrate that differential expectancy effects would surface and be detectable from extremely brief samples of context-minimal teacher behaviour.

The data in this study consisted of judges' ratings of extremely brief (10-second) clips of teachers' verbal and non-verbal behaviour in separate channels (face, body, audio, transcript) when *talking about* and *talking to high expectancy and low expectancy students*. The major question was whether teachers' expectancy effects (manifested in differential behaviour toward high and low expectancy students) could be diagnosed from very brief, relatively context-free, samples of their

behaviour. The investigated teachers were further divided into two groups according to their susceptibility to biasing information, and a second research question focused on the relationship between bias susceptibility and the intensity of expectancy effects.

Although the phenomenon of teachers' self-fulfilling prophecies is well documented, commonly known and generally accepted (Dusek, 1985, p. 2), there is an ongoing controversy over the magnitude and meaning of teacher expectancy effects. Some researchers argue that expectancies have effects of substantial magnitude (e.g. Rosenthal and Rubin, 1978; Smith, 1980; Babad *et al.*, 1982a; Harris and Rosenthal, 1985; Rosenthal, 1985; Babad, 1988), while others argue that teacher expectancy effects are negligible, that differential expectancy behaviour primarily reflects real ability differences among students, and that teachers' differential behaviour is "appropriate, reality-based, and open to corrective feedback" (Brophy, 1983, 1985; see also Hall and Merkel, 1985). With regard to the specific behaviours which mediate expectancy effects, Harris and Rosenthal (1985) conducted meta analyses on 31 behaviours derived from 135 mediation studies and clustered them into a four-factor model of mediation (climate, feedback, input, and output), while Brophy (1983, 1985) summarised 17 different behaviour categories mediating teacher expectancies. Several theoretical models have been proposed to conceptualise the mediation of teacher expectancies (Rosenthal, 1973, 1985; Cooper, 1979, 1985; Brophy, 1983, 1985; Raudenbush, 1984; Peterson and Barger, 1985).

Harris and Rosenthal (1985) distinguished between the link connecting expectations with teacher behaviour and the "self-fulfilling" link connecting differential teacher behaviour with student performance. The present research is focused on the former link only. The term "expectancy effect", as used in this report, refers to differential teacher behaviour associated with varying teacher expectations. Only naturally existing expectations toward teacher-nominated high and low expectancy students were investigated in this study. Thus, the present "expectancy effect" should be distinguished from experimentally manipulated self-fulfilling prophecy effects based on bogus information.

Research on the mediation of teacher expectancies (Dusek, 1985; Harris and Rosenthal, 1985) repeatedly indicates the subtlety and elusive nature of influential teacher behaviours transmitted to students. Very fine nuances in teacher behaviour — many of which are non-verbal, uncontrollable, and often undetected in natural observation — might have substantial, accumulating effects on students.

In studying teacher behaviour, educational researchers have traditionally used context-dependent measurement, essentially based on natural observations in the classroom. Investigators of non-verbal communication (e.g., Rosenthal, 1979; Rosenthal *et al.*, 1979) employed an opposite strategy — separating channels and eliminating most of the content — so as to be able to ascribe raters' judgments to particular channels and combinations of channels. In a recent study, for example, O'Sullivan *et al.*, (1985) predicted judgments based on combined channels (i.e., when a fuller context is available) from judgments based on separate channels such as tone of voice, face alone, or body alone.

Two components of "context" were considered in developing the context-minimal method of measuring teacher behaviour which is utilised in the present investigations: (1) length of judges' exposure to teacher behaviour; and (2) separation of channels. In a natural classroom observation, exposure is very long, allowing the observer to follow events throughout their entire course, thus obscuring

(and potentially biasing) the judgment of specific segments of behaviour. In addition, the observer is exposed to multi-channel information — visual and auditory, verbal and non-verbal. The use of VTR makes it possible to focus on the teacher alone without exposure to the students; to separate different verbal and non-verbal channels in teacher behaviour; and to control the length of judges' exposure to segments of teacher behaviour.

Another distinction employed in this study is that between talking *about* students and talking *to* students. In a recent study, Rosenthal *et al.* (1984) measured therapists' tone of voice when talking about their patients and when talking directly to them. They found that very brief (20-second) clips of content-filtered speech when talking about patients were predictive of therapists' tone of voice in actual interactions with their patients.

The questions presented thus far suggest a within-teacher design. In the selection of teachers for this study, a between-teacher dimension was added: susceptibility to biasing information. Babad (1979) developed a method of identifying susceptibility to biasing information, based on scoring of drawings allegedly made by a high and a low status child. This method was used in a series of studies over the last decade, and the results confirmed the validity of this dimension: susceptibility to bias was found related to extremity of held attitudes (Babad, 1979, 1985), to the tendency to show halo effects (Babad *et al.*, 1982b), and to several other attributes characteristic of the dogmatic cognitive style (Babad and Inbar, 1981; Babad, 1988); biased teachers differed substantially from unbiased teachers in their teaching style (Babad and Inbar, 1981); self-fulfilling prophecy effects in teacher behaviour and student performance (mostly Golem effects) were found for biased, but not for unbiased teachers (Babad *et al.*, 1982a); biased (but not unbiased) teachers were found to show leakage of uncontrollable negative affect in their non-verbal behaviour (Babad *et al.*, 1989); and a longitudinal follow-up study traced a variety of systematic differences between biased and unbiased teachers-in-training in intellectual ability, academic performance, field work evaluations, and drop-out rates (Babad, 1989).

METHOD

Teachers

A group of 123 experienced preschool and elementary school female teachers, in in-service training for a "senior teacher" certificate in an Israeli teacher training college, served as the initial group from which the present sample was drawn. Babad's (1979) method of measuring susceptibility to biasing information was used to identify biased and unbiased teachers. The teachers were taught the scoring procedure of the Goodenough-Harris Draw-A-Person Test (Harris, 1963) and were then asked — under the guise of a reliability exercise — to score two drawings (actually reproduced from the test manual) allegedly drawn by a high status and a low status child. High and low status were created by fictitious information — including name (European or Moroccan) and SES information (parents' education and occupation). The difference between the scores attributed to the two drawings identified subjects' level of susceptibility to biasing information.

A sample of 21 teachers was selected from the large group, representing extreme groups of biased and unbiased teachers. The group of biased teachers, drawn from the extreme end of the distribution, consisted of 14 teachers, and the unbiased group consisted of seven teachers. (The objective, test manual difference between the two drawings is three points in favour of the drawing attributed to the high status child. Thus, a 3-point difference is considered as "no bias". The unbiased group had a mean difference of 2.86 points between the two drawings, the

biased group had a mean difference of 10.36 points, whereas the overall mean difference for the entire group of 123 teachers was 4.8 points.)

All 21 teachers in the sample were female, married or widowed, and mothers of children. Their ages ranged from 27 to 55 years, with a mean and median age in mid- to late-30s. Their teaching experience ranged from five to 31 years, with a mean and median between 15 and 17 years. The two groups of teachers did not differ from each other in any of these demographic characteristics. Of the 21 teachers, nine were preschool teachers, five were elementary school teachers, and seven were remedial teachers.

With a sample of teachers as small as this, sample representativeness could not be ascertained. However, all teachers in this sample were accepted for a training programme leading to a "senior teacher" status. They were sufficiently experienced, passed selective screening, and were positively evaluated, and recommended, by their supervisors. Thus, the possibility of a negative sampling bias must be ruled out.

All 21 teachers were videotaped in their classrooms in the initial warm-up stage. Owing to various uncontrollable circumstances, some teachers were not able to complete all stages of data collection: because of technical VTR problems, the "talking about" clips of two teachers could not be utilised (leaving the "about" sample at 19 teachers — 12 biased and seven unbiased), and because of the shortness of the break and unavailability of some students, only 11 (six biased and five unbiased) teachers completed the subsequent "talking to" stage. The small size of the videotaped sample is regrettable. The authors were aware of the price they would have to pay in terms of levels of significance, but assumed that estimates of effect magnitudes would be relatively uninfluenced by the sample size, making it possible to compare the present results with those previously reported.

The judges were 15 advanced undergraduate students in educational psychology at the Hebrew University of Jerusalem. All judges were Israeli females in their early to mid-20s. The judges were paid to rate the clips, and they were not informed about the intent of the study, the experimental conditions or the stratification of teachers in the sample. (In retrospect, it is regrettable that college students were used as judges. It would have been useful also to employ elementary school children and teachers as judges. Steps are taken to rectify this drawback in current research.)

Videotaping in the classroom

All teachers were first videotaped in their classrooms addressing their entire classes (see Babad *et al.*, 1987, 1988). The camera was positioned in the back of the classroom, and teachers could not detect whether it was filming or not at a given moment.

When the class session was over, the teacher remained in the classroom, facing the camera. Each teacher was asked to choose two children — a good student of high potential, and a weak student of poor potential — and briefly (2-3 minutes) describe each child to the experimenter. Three 10-second clips were recorded in the middle of each description, consisting (in random order) of the face alone, the body alone (from the neck down), and face+body (i.e., the "whole" person). The teacher's speech was recorded in all clips, and one was chosen at random to be used for the audio and transcript clips. All nine clips, representing the various channels and combinations of channels, were later derived from these three 10-second segments. These clips constituted the "talking about" condition.

For the subsequent "talking to" condition, the described students were individually summoned to the classroom, to be taught briefly (2-5 minutes) by the teacher. Some topics were suggested (e.g., north-south, temperatures, speed, etc.) but each teacher was free to teach anything she chose. No particular trend in selection of teaching topics was detected. Three 10-second clips were recorded in the middle of each teaching session, consisting again of the face, the body, and the entire person (plus audio in all recordings). Care was taken that the child would not appear in any of the pictures.

Preparation of clips

Altogether, 36 clips were prepared for each teacher, nine each for the four conditions in the 2×2 design. The nine channels were: (1) *Face only* (no speech); (2) *Body only* (no speech); (3) *Face+Body* (no speech); (4) *Audio only* (recorded speech, no video); (5) *Face+Audio*; (6) *Body+Audio*; (7) *Face+Body+Audio*; (8) *Transcript* (a written account of the words spoken in the 10-second segment, no video and no audio); (9) *Content-filtered speech* (a process that removes from the tape the high frequencies on which word recognition depends but which preserves sequence and rhythm — see Rogers *et al.*, 1971; Rosenthal *et al.*, 1984).

All 36 clips for each teacher were set in a fixed randomised order, to be viewed and rated by the 15 judges. The ninth channel, content-filtered speech, was later dropped from the analyses, since very loud background noises of recess in the schools blended with, and obscured, the teachers' voices.

Four combinations of channels were later added for data analysis, averaging judges' ratings. They were: (1) *Video only* (combination of face, body, and face+body); (2) *Verbal present* (combination of transcript, audio, face+audio, body+audio, and face+body+audio); (3) *Face present* (combination of face, face+body, face+audio, and face+body+audio); and (4) *All eight channels combined* (i.e., the grand mean across clips).

Judges' ratings

The 15 judges viewed all 36 clips of each teacher, and rated each clip on a series of 9-point scales. The scales were: (1) *Warm*; (2) *Dominant*; (3) *Task oriented*; (4) *Tense/Nervous/Anxious*; (5) *Condescending*; (6) *Hostile*; (7) *Clear in communication*; (8) *Democratic*; (9) *Active/energetic/enthusiastic*; and (10) *Flexible*. These rating scales represent typical behaviours measured in mediation of expectancy research (Harris and Rosenthal, 1985). The scales were not further described or explained to the judges in any fashion, but none of the judges expressed any difficulty in understanding these scales.

The judges were not given any information about the purpose of the research, its design, or the specific conditions under which the clips were collected. They were only told that the research dealt with verbal and non-verbal elements of teacher-student interaction in the classroom.

Clips were randomised within teachers, and the judges rated all 36 clips for a given teacher before moving on to the next set of 36 clips. The rationale for this decision was based on the assumption that each teacher was a "constant", and the purpose was to find differentiations within her behaviour. If all clips of all teachers were mixed, gross between-teacher effects would have emerged, swamping the fine within-teacher differences. The method chosen was reasoned to increase the judges' precision in making finer distinctions among the clips for each teacher.

Principal components analysis and composite scores

The means of the judges' ratings in the four experimental conditions, with each channel taken as a separate observation for each teacher, were correlated, and a principal components analysis was computed. This analysis yielded three clear and interpretable factors after varimax rotation: (1) *Non-Dogmatic Behaviour* (consisting of flexible, democratic, and warm — with factor loadings of -0.89 to -0.91); (2) *Negative Affect* (consisting of hostile, condescending, and tense/nervous/anxious — with factor loadings of 0.77 to 0.95); and (3) *Active Teaching Behaviour* (consisting of dominant, active/energetic/enthusiastic, clear, and task-oriented — with factor loadings of 0.63 to 0.87). On the basis of these results, three composite variables were created by averaging the relevant ratings, and all subsequent results are reported for these composite variables only. Prior to combining variables to create composite variables, SDs were checked for homogeneity, so that ingredients of the composite were given essentially similar weights (Rosenthal, 1982).

Correlations between the three composite variables were computed for each channel and for all channels combined. Dogmatic behaviour (the non-dogmatic score reversed) and negative affect were closely related (range of 0.57 to 0.90 , and median correlation of 0.81). Active teaching behaviour was negatively related to both affective variables (a range of -0.27 to -0.72 and a median correlation of -0.53 for dogmatic behaviour; and a range of 0.17 to -0.94 , and a median correlation of -0.38 for negative affect).

Reliability of judges' ratings

Split-half reliabilities were computed by correlating the mean ratings made by the first seven judges with the mean ratings made by the remaining eight judges. This type of correlation estimates the effective reliability of any randomly chosen group of eight judges with any other randomly chosen group of seven. The Spearman-Brown formula for correcting reliabilities to reflect the number of raters used (Rosenthal, 1982) was then applied. The reliabilities reported below are based upon the 15 judges used in the present study. Reliabilities of the judges' ratings were computed separately for each channel and experimental condition, and it was found that they did not vary substantially across condition or channel. Effective reliabilities ranged from 0.54 (when rating transcripts for negative affect) to 0.94 (when rating the face for dogmatism), with a median of 0.82 . Median correlations for the eight channels across the three rating variables ranged from 0.74 (for Body+Audio) to 0.90 (for Face+Body). Across the eight channels, active teaching was the most reliably rated (median $r=0.88$), followed by dogmatic behaviour (median $r=0.85$), and negative affect (median $r=0.75$).

RESULTS

The basic mode of data analysis was a 2×2 analysis of variance with the bias level of the teacher (biased versus unbiased) a between-teachers factor and the expectancy level of the student (high versus low expectancy) a within-teachers factor. Since in any 2^k factorial all main effects and all interactions are single df comparisons or contrasts, they can be tested earlier by the F of the anova or by the t -test of the contrast, since for any contrast $t = \sqrt{F}$ (Rosenthal and Rosnow, 1985). Because t -tests are more useful than F -tests in subsequent statistical procedures (because of their directionality) and because t -tests permit one- or two-

tailed testing of the specific research hypothesis, *ts* are reported along with their associated one-tailed *P* values (Rosenthal, 1984). Should readers desire any of the *ts* in the *F* form, the *ts* need only be squared. Should any of the one-tailed tests be desired as two-tailed, of course, they need only be doubled. A separate 2×2 analysis of variance (or set of three orthogonal contrasts — bias effect, expectancy effect, and interaction effect) was computed for each channel, combination of channels, and for each of the composite variables. Means of the composite scores for all channels in both conditions are presented in Table 1.

TABLE 1
MEANS OF JUDGES' RATINGS OF THE THREE COMPOSITE SCORES FOR ALL CHANNELS IN THE TALKING ABOUT AND TALKING TO CONDITIONS

CHANNELS	DOGMATIC BEHAVIOUR		NEGATIVE AFFECT		ACTIVE TEACHING	
	HIGH EXP.	LOW EXP.	HIGH EXP.	LOW EXP.	HIGH EXP.	LOW EXP.
ABOUT STUDENTS (<i>N</i> = 19)						
Face	4.7	5.0	3.3	3.4	5.9	5.8
Body	5.4	5.3	3.6	3.7	4.9	4.9
Face+Body	4.9	5.0	3.3	3.5	5.7	5.6
Transcript	4.7	5.2	3.6	3.7	6.0	5.8
Audio	5.0	5.1	3.6	3.5	5.9	5.8
Face+Audio	4.7	4.9	3.5	3.6	6.1	6.2
Body+Audio	4.9	5.1	3.4	3.5	6.1	5.9
Face+Body+Audio	4.8	4.8	3.3	3.5	5.9	6.1
TO STUDENTS (<i>N</i> = 11)						
Face	4.8	5.0	3.1	3.5	5.9	6.4
Body	5.0	5.0	3.5	3.5	6.1	6.0
Face+Body	4.7	4.6	3.3	3.3	6.2	6.5
Transcript	4.5	4.4	3.4	3.5	6.4	6.5
Audio	4.9	4.6	3.5	3.5	5.9	5.8
Face+Audio	4.4	4.6	3.4	3.5	6.5	6.6
Body+Audio	4.7	4.5	3.4	3.4	6.6	6.6
Face+Body+Audio	4.5	4.5	3.4	3.5	6.6	6.7

Expectancy effects for "to" and "about" combined

Initial analyses are reported to demonstrate the existence of overall expectancy effects of substantial magnitude in context-free measurement. The results in Table 2 were computed for *all* clips, combining the talking about and talking to conditions.

Expectancy effects were found mostly for ratings of negative affect — teachers were judged to show more negative affect when the low expectancy students were concerned. This was evident in the combination of all eight channels ($t = 2.15^{**}$, $r = 0.58$), but was contributed to especially by the verbal material ($t = 2.22^{**}$, $r = 0.60$) and particularly by the transcript, that is, the actual words spoken ($t = 2.40^{**}$, $r = 0.62$). Thus, more negative words were spoken about and to low expectancy students as compared with high expectancy students. Expectancy

effects of substantial magnitude were found for the face present channel as well. Thus, the results in Table 2 confirmed that teacher expectancies can be diagnosed from brief, context-minimal exposure to their verbal and non-verbal behaviour.

TABLE 2

MAIN EFFECTS OF TEACHERS' EXPECTATIONS ON JUDGES' MEAN RATINGS OF 11 TEACHERS TALKING BOTH TO AND ABOUT THEIR STUDENTS

	DOGMATIC BEHAVIOUR (a)		NEGATIVE AFFECT		ACTIVE TEACHING	
	<i>t</i>	<i>r</i>	<i>t</i>	<i>r</i>	<i>t</i>	<i>r</i>
GROUPED CHANNELS						
All Eight	1.39*	0.42	2.15**	0.58	1.30	0.40
Face Present	1.69*	0.49	2.20**	0.59	2.22**	0.59
Verbal Present	1.38*	0.42	2.22**	0.60	0.73	0.24
Video Only	0.69	0.22	1.59*	0.47	1.81*	0.52
INDIVIDUAL CHANNELS						
<i>A. Video Only</i>						
Face	1.57*	0.48	1.65*	0.39	1.96**	0.55
Body	0.00	0.00	1.10	0.34	0.14	-0.05
Face+Body	0.00	0.02	0.30	0.10	1.48*	0.44
<i>B. Verbal Present</i>						
Transcript	1.64*	0.48	2.40**	0.62	0.28	0.09
Audio	0.00	-0.01	0.90	-0.29	1.25	-0.39
Face+Audio	1.30*	0.40	1.46*	0.44	0.99	0.31
Body+Audio	0.01	-0.03	1.74*	0.50	0.00	-0.02
Face+Body+Audio	0.57	0.19	1.84**	0.52	1.53*	0.46

(a) These *t*'s are from ANOVA *F*'s, where $t = \sqrt{F}$

* $P < 0.10$; ** $P < 0.05$; *** $P < 0.01$; all one-tailed

Note: High values of *r* indicate that more of the rating variable was directed at the low-expectancy student.

Talking about students: expectancy and bias effects

Expectancy main effects for teachers talking about their students are presented in Table 3. A very strong expectancy effect was found for dogmatic behaviour ($t = 3.44^{***}$, $r = 0.64$), indicating that the judges detected less warmth, flexibility and democracy when the teachers talked about their low expectancy students. This was found not only in *what* was said (verbal present), but in *how* it was said as well (video only) — both significant at the 0.01 level. Thus, the teachers looked more dogmatic to the judges, especially in their face (face present, $t = 3.03^{***}$, $r = 0.58$), and their words in these 10-second segments were judged more dogmatic when they talked about their low expectancy students (verbal present, $t = 2.82^{***}$, $r = 0.57$).

A strong expectancy effect was found for negative affect ($t = 2.60^{***}$, $r = 0.53$), indicating that the teachers were judged to show more negative affect when talking about their low expectancy students. But, whereas the dogmatic effect was found in both verbal and non-verbal channels, this effect was mostly visual, appearing in the face present channels.

Not surprisingly, no expectancy effects were found for the active teaching variable in Table 3, since the teachers were talking *about* the students rather than actually teaching them.

TABLE 3
MAIN EFFECTS OF TEACHERS' EXPECTATIONS ON JUDGES' MEAN RATINGS OF
19 TEACHERS TALKING ABOUT THEIR STUDENTS

	DOGMATIC BEHAVIOUR (a)		NEGATIVE AFFECT		ACTIVE TEACHING	
	<i>t</i>	<i>r</i>	<i>t</i>	<i>r</i>	<i>t</i>	<i>r</i>
GROUPED CHANNELS						
All Eight	3.44***	0.64	2.60***	0.53	0.42	-0.10
Face Present	3.03***	0.58	2.74***	0.55	0.46	0.11
Verbal Present	2.82***	0.57	1.77**	0.40	0.37	-0.09
Video Only	2.50**	0.52	1.67*	0.38	0.48	-0.11
INDIVIDUAL CHANNELS						
<i>A. Video Only</i>						
Face	1.78**	0.40	0.75	0.18	0.59	-0.14
Body	0.17	0.04	1.41*	0.32	0.64	-0.15
Face+Body	1.05	0.25	1.80**	0.41	0.37	-0.09
<i>B. Verbal Present</i>						
Transcript	2.73***	0.55	0.94	0.22	1.02	-0.24
Audio	0.87	0.21	1.27	-0.29	1.72*	-0.39
Face+Audio	2.19**	0.47	1.36*	0.32	0.41	0.10
Body+Audio	1.96**	0.43	0.93	0.22	0.85	-0.20
Face+Body+Audio	0.30	0.07	2.60***	0.53	1.79**	0.40

(a) These *t*'s are from ANOVA *F*'s, where $t = \sqrt{F}$

* $P < 0.10$; ** $P < 0.05$; *** $P < 0.01$; all one-tailed

Note: High values of *r* indicate that more of the rating variable was shown when talking about the low-expectancy student.

Main effects associated with the teachers' bias type are presented in Table 4. Biased teachers differed from unbiased teachers in their style of talking about their students. The differences were clustered in ratings of negative affect, indicating that biased teachers showed more negative affect than unbiased teachers when talking about their students. Differential negative affect was detected both in the face ($t = 2.17^{**}$, $r = 0.47$) and in what was said (transcript: $t = 2.03^{**}$, $r = 0.44$; and face+audio: $t = 2.00^{**}$, $r = 0.44$).

Main effects of teachers' bias in ratings of transcript were significant for all three composite variables (dogmatic behaviour: $t = 2.34^{**}$, $r = 0.49$; negative affect: $t = 2.03^{**}$, $r = 0.44$; active teaching behaviour: $t = 1.81^{**}$, $r = 0.40$). Thus, in terms of the actual words spoken about students in the 10-second clips, biased teachers were judged as less democratic, flexible, and warm, and as more hostile, condescending and tense/nervous/anxious than unbiased teachers, whereas the words spoken by unbiased teachers were rated higher on clarity, dominance, task orientation, and activity/energy/enthusiasm.

TABLE 4

MAIN EFFECTS OF TEACHERS' BIAS TYPE ON JUDGES' MEAN RATINGS OF
19 TEACHERS TALKING ABOUT THEIR STUDENTS

	DOGMATIC BEHAVIOUR (a)		NEGATIVE AFFECT		ACTIVE TEACHING	
	<i>t</i>	<i>r</i>	<i>t</i>	<i>r</i>	<i>t</i>	<i>r</i>
GROUPED CHANNELS						
All Eight	0.35	0.08	1.29	0.30	0.10	0.02
Face Present	0.35	0.08	1.72*	0.38	0.17	-0.04
Verbal Present	0.14	0.03	1.52*	0.35	0.51	0.12
Video Only	0.60	0.14	0.65	0.16	0.45	-0.11
INDIVIDUAL CHANNELS						
<i>A · Video Only</i>						
Face	0.99	0.23	2.17**	0.47	0.17	0.04
Body	0.17	0.04	1.74**	-0.40	0.68	-0.17
Face+Body	0.00	0.00	0.74	0.18	0.17	-0.04
<i>B · Verbal Present</i>						
Transcript	2.34**	0.49	2.03**	0.44	1.81**	-0.40
Audio	0.47	0.11	0.14	0.03	0.57	0.14
Face+Audio	0.72	0.17	2.00**	0.44	0.26	0.06
Body+Audio	1.46*	-0.34	1.35*	0.32	1.45*	0.34
Face+Body+Audio	0.76	-0.18	0.93	0.22	0.84	0.20

(a) These *t*'s are from ANOVA *F*'s, where $t = \sqrt{F}$

* $P < 0.10$; ** $P < 0.05$; *** $P < 0.01$; all one-tailed

Note: High values of *r* indicate that more of the rating variable was shown by high-bias teachers.

Talking to students — expectancy and bias effects

Expectancy effects in talking to students are presented in Table 5. Whereas the expectancy effects in talking about students were concentrated in ratings of dogmatic behaviour and negative affect, the expectancy effects here were found in ratings of negative affect and active teaching behaviour. Teachers were judged to show more negative affect in their facial expressions to the low expectancy student ($t = 2.94^{***}$, $r = 0.70$). It is interesting to note that although the teachers' faces looked more negative to the judges when talking to the low expectancy student, no expectancy differences were picked up from their words. A similar pattern was found for ratings of active teaching behaviour. Teachers were rated to be more active toward the low expectancy student, but again, this was picked up in the face (face present: $t = 2.26^{**}$, $r = 0.60$; video only: $t = 2.09^{**}$, $r = 0.57$; face: $t = 2.09^{**}$, $r = 0.57$), whereas no differences were found for the audio and verbal channels.

Thus expectancy effects in talking to students were found in non-verbal (mostly face) ratings of teachers' negative affect and level of activity and dominance. The expectancy effect in negative affect for all eight channels combined ($t = 1.57^*$, $r = 0.46$) was of substantial magnitude, even though the effect did not reach statistical significance due to the small *N*. Harris and Rosenthal (1985) reported substantial effect magnitudes in meta analyses for the "climate" factor (0.20 to 0.36, see p. 377), based on context-dependent measurement of teachers' behaviour. In terms of the Binomial Effect Size Display (Rosenthal and Rubin, 1982), an *r* of 0.46 would be equivalent to increasing the success rate of a new treatment from 27 per cent to 73 per cent.

TABLE 5

MAIN EFFECTS OF TEACHERS' EXPECTATIONS ON JUDGES' MEAN RATINGS OF
11 TEACHERS TALKING DIRECTLY TO THEIR STUDENTS

	DOGMATIC BEHAVIOUR (a)		NEGATIVE AFFECT		ACTIVE TEACHING	
	<i>t</i>	<i>r</i>	<i>t</i>	<i>r</i>	<i>t</i>	<i>r</i>
GROUPED CHANNELS						
All Eight	0.22	-0.06	1.57*	0.46	0.95	0.23
Face Present	0.48	0.23	1.69*	0.49	2.26**	0.60
Verbal Present	0.42	-0.14	0.96	0.30	0.10	-0.04
Video Only	0.10	0.03	1.88**	0.53	2.09**	0.57
INDIVIDUAL CHANNELS						
<i>A · Video Only</i>						
Face	1.03	0.32	2.94***	0.70	2.09**	0.57
Body	0.17	0.06	0.84	0.27	0.20	-0.07
Face+Body	0.95	-0.30	0.64	-0.21	1.80*	0.51
<i>B · Verbal Present</i>						
Transcript	0.69	-0.22	0.94	0.30	0.14	0.05
Audio	0.98	-0.31	0.30	-0.10	1.01	-0.32
Face+Audio	1.07	0.34	0.90	0.29	0.39	0.13
Body+Audio	0.94	-0.30	0.96	0.31	0.22	-0.07
Face+Body+Audio	0.10	0.04	0.84	0.27	0.26	0.09

(a) These *t*'s are from ANOVA *F*'s, where $t = \sqrt{F}$ * $P < 0.10$; ** $P < 0.05$; *** $P < 0.01$; all one-tailedNote: High values of *r* indicate that more of the rating variable was shown when talking to the low-expectancy student.

The findings on active teaching behaviour are quite interesting. The literature provides two trends in teachers' behaviour toward low expectancy students (see Babad, 1988). In terms of teachers' "input," Harris and Rosenthal (1985) reported that teachers attempt to teach more material and more difficult material to high expectancy students (highly significant *Z*-values, and effect magnitudes around 0.30 in meta-analyses). On the other hand, Brophy (1985), and Hall and Merkel (1985) highlighted the phenomenon of teachers often investing extra efforts in their low expectancy students and pushing them harder, in an attempt (so it would seem) to "compensate" for their disadvantage. And indeed Harris and Rosenthal found in meta-analyses that teachers lecture more, give more directions, use more direct influence, and have more work-related contacts with their low expectancy students. In other words, in certain variables reflecting teachers' direct control, low expectancy students receive more than high expectancy students. This is, in essence, what was found here for the active teaching composite.

It seems that the teachers' faces carried a dual message when communicating to their low expectancy students: task-orientation, clarity, and energetic dominance on the one hand — negative affect and hostility on the other hand. The teachers seemed to push their low students, at the same time transmitting negative messages to them. Maybe they wished to demonstrate that the low expectancy students were not so "bad" after all, pushed them harder, but were also disappointed with them. One might consider for a moment the setting in which the teacher was being videotaped interacting with a student she had just nominated as one of her

"poor" students: teachers were probably trying to act "appropriately" or compensate these students. Nastiness, however, seems to have slipped out. That kind of situation is labelled "leakage" in the deception literature (Ekman and Friesen, 1969, 1974; Rosenthal and DePaulo, 1979; Zuckerman *et al.*, 1986) — the situation where more positive affect is (deceitfully) transmitted in more controllable channels (e.g., transcript, followed by face), while negative affect is given away through less controllable, "leakier" channels such as the body. (See Babad *et al.*, 1989, on leakage in the behaviour of the biased teachers.)

ANOVAs of bias effects in talking to all students yielded no significant differences between biased and unbiased teachers, and only one effect (dogmatic behaviour in the body) reached the 0.05 level of significance. However, some of the effect magnitudes were substantial, indicating a clear trend for biased teachers to be more dogmatic and somewhat less active and dominant than unbiased teachers. For dogmatic behaviour, the effect magnitude for all eight channels combined was 0.35, contributed by both non-verbal and verbal aspects (video only: $r = 0.36$; face present: $r = 0.30$; face+body: $r = 0.30$; transcript: $r = 0.37$; face+body+audio: $r = 0.33$; audio: $r = 0.30$; verbal present: $r = 0.30$). Babad *et al.*, (1982a) reported an effect magnitude of 0.65 for bias in dogmatic behaviour in context-dependent measurement. The effects reported here for dogmatic behaviour in context-minimal measurement are smaller, yet substantial: in terms of the Binomial Effect Size Display (Rosenthal and Rubin, 1982), an r of 0.35 would be equivalent to increasing the success rate of a new treatment from 33 per cent to 67 per cent.

The effects of bias in active teaching behaviour were somewhat smaller (effect magnitude for all eight channels: $r = 0.26$), but they were also indicated in both non-verbal and verbal aspects (video only: $r = 0.32$; body: $r = 0.44$; face+body: $r = 0.25$; transcript: $r = 0.28$; verbal present: $r = 0.19$). In all cases, biased teachers were rated as less dominant, task-oriented, clear, and active/energetic/enthusiastic than unbiased teachers when talking to students.

Several expectancy $\times$ bias interaction effects were found in the "talking to" ANOVAs. They were mostly concentrated in the verbal aspects of the dogmatic behaviour ratings. Biased teachers showed a stronger expectancy effect than unbiased teachers, that is, showed a greater differential in dogmatic verbal behaviour by being more dogmatic than unbiased teachers to the low expectancy students. The most notable effects were found for transcript ($t = 2.44^{**}$, $r = 0.63$), face+audio ($t = 1.90^{**}$, $r = 0.53$) and the verbal present combination ($t = 1.86^{**}$, $r = 0.53$).

DISCUSSION

The analysis of teacher behaviour via a context-minimal method, in which judges rate extremely brief samples of behaviour broken into separate non-verbal and verbal channels, yielded clear and systematic expectancy effects, their magnitude equalling expectancy effects found in full context observations (Harris and Rosenthal, 1985). Teachers' views of their high and low students permeated the most minute and molecular elements of their behaviour and were detected by the judges. Moreover, the data were obtained when the teachers were facing the camera, and they might well have attempted to put on their best behaviour. Expectancy effects might thus be more pervasive and deep-seated than some researchers and educators might wish to believe.

Negative affect was transmitted to low expectancy students mostly through the non-verbal channels, particularly the face. If judges could pick up these negative feelings from isolated 10-second clips when teachers might have been on their best behaviour, the accumulation of negative affective messages absorbed by low expectancy students over years of continuous interaction must be rather intense!

While the teachers were more negative emotionally when talking about and to their low expectancy students, they compensated by directing more active teaching behaviour at them. The compensation and special investment in low ability students might be the factor persuading educators (and investigators) to believe that teacher expectancies are not harmful. But such compensation does not nullify the negative non-verbal communication which was picked up so readily by the judges. In fact, over-compensation might even come to be perceived by students as a symbolic signal of low expectancy. . . .

Expectancy effects were found in both talking about and talking to conditions. This is in line with Rosenthal *et al.*, (1984), who showed that therapists' tone of voice when talking about their patients could meaningfully predict their tone of voice when they actually talked to their patients.

In terms of the communication process, the "about" and "to" conditions were characterised by different patterns. The factor common to both conditions was that more negative affect was transmitted to the low expectancy students. In the "about" condition, that was accompanied by more dogmatic behaviour directed at low expectancy students in non-verbal and verbal channels, while in the "to" condition, low expectancy students were compensated by more active teaching activity in the non-verbal channels.

This difference can probably be explained in terms of teachers' attempts at self control. They might have felt freer to express themselves openly when talking about students, but more restrained and controlled in actual interaction with students. Since it is easier to control one's words than one's expressions, expectancy effects in the verbal channels were found only in the "about", but not in the "to" condition. The non-verbal negative expressions in the "to" condition were probably manifestations of leakage. Conceptually, it makes perfect sense that intentional compensation of low expectancy students will be accompanied by non-verbal leakage in less controllable channels. The authors (Babad *et al.*, 1989) analysed non-verbal leakage of these teachers in videotaped segments when they were addressing their entire classes. They found that biased teachers showed substantial leakage in both dogmatic behaviour and negative affect, whereas unbiased teachers showed no leakage while addressing their entire classes.

The comparison among channels leads to an interesting conclusion which justifies, in retrospect, the separation of channels in the present context-minimal measurement. The Face+Body+Audio channel supposedly contains more information than any other individual channel employed, and, except for the brevity of exposure, is the closest to more traditional methods of observation. And yet the results for the F+B+A channel in all tables are weaker and less conclusive than the results found for some other channels (e.g., face) which are supposedly less informative. This is probably due to the fact that different channels provided different types of information which were sometimes dissonant, or even contradictory.

With regard to the susceptibility to bias dimension, several main effects and interaction effects reported above indicated either that biased teachers were more

negative in their communication style, or that they manifested stronger expectancy effects than unbiased teachers. These findings indicate that the division of teachers according to their level of susceptibility to biasing information made a difference, and biased teachers differed from unbiased teachers in some aspects of their behaviour, in line with previous reports. On the other hand, the bias dimension did not have the same power and intensity in predicting teacher behaviour as reported in the previous studies (Babad and Inbar, 1981; Babad *et al.*, 1982a, 1982b; Babad *et al.*, 1988). Several reasons might account for this gap: First, Babad *et al.* (1982a) found expectancy effects (and Babad *et al.*, 1989, found leakage effects) only for biased teachers, but not for unbiased teachers. In the present study, expectancy effects were found for *all* (biased and unbiased) teachers. What this might mean is that, when the smallest units of behaviour and the most isolated elements of communication are analysed in a context-free evaluation, even teachers who are relatively unsusceptible to stereotypically biasing information might show expectancy differentials in their behaviour. A complementary explanation might be that the bias dimension (as defined and measured) is a stronger predictor in longer, context-dependent observations. Finally, this study was methodologically geared to pick up within-teacher differences, whereas bias effects represent between-teacher differences. The judges viewed all 36 clips of one teacher before they moved on to the next teacher. This method tended to sharpen the judges' distinctions between clips, thereby maximising expectancy effects and condition effects and minimising between-teacher differences.

Correspondence and requests for reprints should be addressed to Professor Elisha Babad, School of Education, Hebrew University of Jerusalem, Mount Scopus, Jerusalem, Israel 91 905.

REFERENCES

- BABAD, E. (1979). Personality correlates of susceptibility to biasing information. *J. Person. soc. Psychol.*, 37, 195-202.
- BABAD, E. (1985). Some correlates of teachers' expectancy bias. *Am. educ. Res. J.*, 22, 175-183.
- BABAD, E. (1988). "Pygmalion" — twenty years after. Unpublished manuscript. The School of Education, Hebrew University of Jerusalem.
- BABAD, E. (1989). A longitudinal follow-up study of biased and unbiased teachers-in-training. Unpublished manuscript. The School of Education, Hebrew University of Jerusalem.
- BABAD, E., and INBAR, J. (1981). Performance and personality correlates of teachers' susceptibility to biasing information. *J. Person. soc. Psychol.*, 40, 553-561.
- BABAD, E., INBAR, J., and ROSENTHAL, R. (1982a). Pygmalion, Galatea, and the Golem: investigations of biased and unbiased teachers. *J. educ. Psychol.*, 74, 459-474.
- BABAD, E., INBAR, J., and ROSENTHAL, R. (1982b). Teachers' judgments of students' potential as a function of their own susceptibility to biasing information. *J. Person. soc. Psychol.*, 42, 541-547.
- BABAD, E., BERNIERI, F., and ROSENTHAL, R. (1987). Nonverbal and verbal behaviour of preschool, remedial, and elementary school teachers. *Am. educ. Res. J.*, 24, 405-415.
- BABAD, E., BERNIERI, F., and ROSENTHAL, R. (1989). Nonverbal communication and leakage in the behaviour of biased and unbiased teachers. *J. Person. soc. Psychol.*, 56, 89-94.
- BLANCK, P. D. (1987). The "process" of field research in the courtroom: a descriptive analysis. *Law hum. Behav.*, 28, 437-442.
- BLANCK, P., ROSENTHAL, R., VANNICELLI, M., and LEE, T. D. (1986). Therapists' tone of voice: descriptive, psychometric, interactional, and competence analysis. *J. soc. clin. Psychol.*, 4, 154-178.
- BROPHY, J. E. (1983). Research on the self-fulfilling prophecy and teacher expectations. *J. educ. Psychol.*, 75, 631-661.
- BROPHY, J. E. (1985). Teacher-student interaction. In DUSEK, J. B. (Ed.), *Teacher Expectancies*. Hillsdale, N.J.: Lawrence Erlbaum.
- COOPER, H. (1979). Pygmalion grows up: a model for teacher expectation communication and performance influence. *Rev. educ. Res.*, 49, 389-410.
- COOPER, H. (1985). Models of teacher expectation communication. In DUSEK, J. B. (Ed.), *Teacher Expectancies*. Hillsdale, N.J.: Lawrence Erlbaum.
- DUSEK, J. B. (Ed.) (1985). *Teacher Expectancies*. Hillsdale, N.J.: Lawrence Erlbaum.
- EKMAN, P., and FRIESEN, W. V. (1969). Nonverbal leakage and clues to deception. *Psychiatry*, 32, 88-106.

- EKMAN, P., and FRIESEN, W. V. (1974). Detecting deception from the body or face. *J. Person. soc. Psychol.*, 29, 288-298.
- HALL, V. C., and MERKEL, S. P. (1985). Teacher expectancy effects and educational psychology. In DUSEK, J. B. (Ed.), *Teacher Expectancies*. Hillsdale, N.J.: Lawrence Erlbaum.
- HARRIS, D. B. (1963). *Goodenough-Harris Drawing Test Manual*. New York: Harcourt, Brace, and World.
- HARRIS, M.J., and ROSENTHAL, R. (1985). Mediation of interpersonal expectancy effects: 31 meta-analyses. *Psychol. Bull.*, 97, 363-386.
- MILMOE, S., ROSENTHAL, R., BLANE, H. T., CHAFETZ, M. E., and WOLF, I. (1967). The doctor's voice: postdictor of successful referral of alcoholic patients. *J. abnorm. Psychol.*, 72, 78-84.
- O'SULLIVAN, M., EKMAN, P., FRIESEN, W., and SCHERER, K. (1985). What you say and how you say it: the contribution of speech content and voice quality to judgments of others. *J. Person. soc. Psychol.*, 48, 54-62.
- PETERSON, P. L., and BARGER, S. A. (1985). Attribution theory and teacher expectancy. In DUSEK, J. B. (Ed.), *Teacher Expectancies*. Hillsdale, N.J.: Lawrence Erlbaum.
- RAUDENBUSH, S. W. (1984). Magnitude of teacher expectancy effects on pupil IQ as a function of the credibility of expectancy induction: a synthesis of findings from 18 experiments. *J. educ. Psychol.*, 76, 85-97.
- ROGERS, P. L., SCHERER, K., and ROSENTHAL, R. (1971). Content- filtering human speech. *Behav. Res. Meth. Instrum.*, 3, 16-18.
- ROSENTHAL, R. (1973). The mediation of Pygmalion effects: a four factor "theory". *Papua New Guinea J. Educ.*, 9, 1-12.
- ROSENTHAL, R. (Ed.) (1979). *Skill in Nonverbal Communication*. Cambridge, Mass.: Oelgeschlager, Gunn, and Hain.
- ROSENTHAL, R. (1982). Conducting judgment studies. In SCHERER, K. R., and ECKMAN, P. (Eds.), *Handbook of Methods in Nonverbal Behaviour Research*. Cambridge: Cambridge University Press.
- ROSENTHAL, R. (1984). *Meta-analytic Procedures for Social Research*. London: Sage.
- ROSENTHAL, R. (1985). From unconscious experimenter bias to teacher expectancy effects. In DUSEK, J. B. (Ed.), *Teacher Expectancies*. Hillsdale, N.J.: Lawrence Erlbaum.
- ROSENTHAL, R., and JACOBSON, L. (1968). *Pygmalion in the Classroom*. New York: Holt, Rinehart, and Winston.
- ROSENTHAL, R., and RUBIN, D. B. (1978). Interpersonal expectancy effects: the first 345 studies. *Behav. Brain Sci.*, 3, 377-415.
- ROSENTHAL, R., and DEPAULO, B. M. (1979). Sex differences in accommodation in nonverbal communication. In ROSENTHAL, R. (Ed.), *Skill in Nonverbal Communication*. Cambridge, Mass.: Oelgeschlager, Gunn, and Hain.
- ROSENTHAL, R., and RUBIN, D. B. (1982). A simple general purpose display of magnitude of experimental effect. *J. educ. Psychol.*, 74, 166-169.
- ROSENTHAL, R., and ROSNOW, R. L. (1985). *Contrast Analysis: Focused Comparisons in the Analysis of Variance*. London: Cambridge University Press.
- ROSENTHAL, R., HALL, J. A., DiMATTEO, M. R., ROGERS, P. L., and ARCHER, D. (1979). *Sensitivity to Nonverbal Communication: The PONS Test*. Baltimore: Johns Hopkins University Press.
- ROSENTHAL, R., BLANCK, P. D., and VANNICELLI, M. (1984). Speaking to and about patients: predicting therapists' tone of voice. *J. consult. clin. Psychol.*, 52, 679-686.
- SMITH, M. L. (1980). Teacher expectations. *Eval. Educ.*, 4, 53-55.
- ZUCKERMAN, M., DEPAULO, B. M., and ROSENTHAL, R. (1986). Humans as deceivers and lie detectors. In BLANCK, P., BUCK, R., and ROSENTHAL, R. (Eds.), *Nonverbal Communication in the Clinical Context*. University Park, Pennsylvania State University Press.

(Manuscript received 6th June, 1988)

DETECTION OF MISSING AND IRRELEVANT INFORMATION WITHIN ALGEBRAIC STORY PROBLEMS

BY RENAE LOW AND RAY OVER
(La Trobe University, Victoria, Australia)

SUMMARY. Schematic knowledge was assessed through text editing of algebraic story problems. Seventy-two students in tenth grade were required to denote what information essential to solution was missing from problems, and to detect irrelevant information within problems. The capacity of students to identify the necessary and sufficient information needed for solution of problems accounted for 90 per cent of variance in the solution rates for these same problems. Further, text editing correlated highly with the ability of students to determine whether algebraic story problems were similar to or different from each other. The results offer support to claims that knowledge of problem structure is necessary for solution of algebraic story problems.

INTRODUCTION

It is claimed in models of mathematical ability that place emphasis on information processing that solution of problems is dependent not just on the exercise of computational skills but on access to domain-specific knowledge. For example, Mayer (1983) has proposed that in order to solve the classic motorboat problem ("The motorboat travelled downstream in 120 minutes with a current of five miles per hour. The return upstream trip against the same current took three hours. What was the speed of the boat in still water?"), a person requires at least five categories of knowledge: linguistic (e.g., what "still water" means); semantic (e.g., that rivers have currents that run only downstream, but boats can travel upstream and downstream); schematic (that river current problem involves specific relationships between speed of the boat, rate of the current, and time); algorithmic (how to perform a given sequence of operations); and strategic (standard solution processes such as moving the unknown to one side of the equation).

Schematic knowledge refers to the understanding of the structure of a problem. A person who is able to identify the problem as belonging to a given type (e.g. river current contrast) is able to determine what information from the text of the problem should be used, in what sequence, and through what operations. In order to solve any version of the classic motorboat problem, the person has to be able to realise that the specific details provided in the problem at hand must be organised into the equation $(\text{rate of boat} + \text{rate of current}) \times (\text{time downstream}) = (\text{rate of boat} - \text{rate of current}) \times (\text{time upstream})$.

Hinsley *et al.* (1977) studied schematic knowledge by requiring adults to group algebra problems into clusters, categorise problems after hearing only part of the text, provide answers to problems when content words were replaced by nonsense words, and solve problems when material in the text was ambiguous. They found people were consistent in grouping algebra word problems into categories. Further, after hearing the words "A river steamer", many participants in the study judged that the problem was going to be one involving river current contrasts. One comment was, "It's going to be one of those river things with upstream, downstream; and still water. You are going to compare times upstream, downstream — or if the time is constant, it will be the distance" (Hinsley *et al.*, 1977, p. 97). Individuals can thus categorise problems without completely formulating

them for solution. Hinsley *et al.* also found that some people interpreted an ambiguous problem in "triangle" terms and others in "distance-rate-time" terms, depending on the information they had attended to. Facts were misread if they were inconsistent with categorisation: for example, one person misread "four minutes" as "four miles", took this as the length of one side of the triangle, and attempted to solve the problem using Pythagoras' theorem. From a number of converging demonstrations Hinsley *et al.* (1977) concluded that the encoding and retrieval of information in the process of solving algebraic story problems is governed by a person's schematic knowledge.

In further research emphasising the role of problem schemata, Mayer (1982) required adults to recall problems to which they had been exposed earlier, and to construct problems within a given theme (e.g., "trains leaving stations"). Recall was best for the types of problems that occur frequently within textbooks, and schemata-relevant material was more likely to be recalled accurately than schemata-irrelevant information. In addition, the problems constructed by students tended to approximate standard rather than non-standard formats. Mayer (1981) has established a framework for classifying algebra story problems. He identified some 20 categories, each with many distinct "templates". For example, motion problems had 13 different templates, including simple distance-rate-time, vehicles approaching from opposite directions, one vehicle overtaking another, and speed changing during a journey. The 2000 algebraic problems that Mayer surveyed involved almost 100 different templates. Mayer argued that errors in solution will result when the text of the problem evokes use of an inappropriate template. He concluded that education in mathematics should involve training in the recognition of categories, as well as templates for major problem types.

There have been several demonstrations that experts and novices differ not only in their ability to produce the correct answer to a problem, but in their mode of cognitive representation of problems (see Larkin *et al.*, 1980; Chi *et al.*, 1981). Novices tend to be influenced by surface characteristics and irrelevant detail rather than by deep structure when categorising problems. In contrast, experts identify similarities and differences between problems in terms of underlying principles. Good and poor novice problem solvers also differ in terms of their cognitive structures (de Jong and Ferguson-Hessler, 1986), and after a period of intensive training in mathematical problem solving novices come to demonstrate schematic knowledge (as indexed by the perception of problem relatedness) much like that of experts (Schoenfeld and Hermann, 1982). In a study of eighth grade students Silver (1979) found that a student's mathematical ability predicted the extent to which a student could classify algebraic problems as related or unrelated. In a further study, Silver (1981) established that students at seventh-grade level who performed well in problem solution were able to recall the structure of a mathematical problem, and not just surface detail, weeks after being presented with the problem. In contrast, poor problem solvers often could remember the context of the problem statement and the question posed in the problem, but typically they had limited recall of structural information.

Most investigators have studied schematic knowledge through reliance on tasks such as memory for problems, problem sorting, and classification. The focus in the present research is on text editing of algebraic story problems. The issue of interest is whether the capacity to determine whether the text of an algebraic story problem contains the necessary and sufficient information for solution is predictive of the ability to solve this problem. There are good reasons for supposing that schematic knowledge is needed for text editing. When presented with an algebraic problem that is incomplete in the sense that further information

must be provided before solution can be reached (e.g., "A rectangular lawn is 72 square metres. What is the width of the lawn?"), a person needs to know that the area of a rectangle is the length of one side multiplied by the length of the adjacent side in order to identify the essential component missing from the problem. Similarly, when presented with a problem such as "Amanda is four years older than Brenda, and Brenda is two years older than David. If the combined ages of Amanda and Brenda amount to 28 years, how old is Amanda?", a person should be able to identify that there is information in the problem that is irrelevant for solution ("and Brenda is two years older than David") only from knowledge of the structure of this class of problem.

The question taken up in the present research is whether text editing skills such as identifying what information is missing from problems and deleting information that is irrelevant are predictive of success in solving problems. If problem solution requires schematic knowledge, a student who can solve a given problem should be able to identify the minimal text that will result in the problem being open to solution. Since calculation skills are also involved, schematic knowledge may be thought of as necessary but not sufficient for solving a problem. In this event, some individuals may be able to provide missing information and to delete irrelevant information on a text editing task, but be unable to arrive at the correct solution to the problem. However, it should never be the case that a person will solve a problem after being unable to edit the text of the problem by specifying missing information or identifying irrelevant information.

In the present study, students in tenth grade were presented with a set of algebraic story problems in five different test formats. In the two formats that entailed text editing, students were required either to denote what information essential to solution was missing from a problem or to detect irrelevant information within a problem. In the third format students were required to solve each problem. The fourth format assessed the students' memory for problems they had previously solved. Recognition memory was measured, with students having to indicate whether or not specific algebraic story problems were ones that had been shown earlier. In the fifth format the four algebraic story problems within a set included three problems that were similar in their schematic knowledge requirements. Students were required to identify the problem within the set that they believed was different from the other three. The aim of the study is to examine relationships between text editing and not only problem solution but also traditional measures of schematic knowledge such as problem categorisation and memory.

METHOD

Sample

Forty-two boys and 34 girls in tenth grade at an independent coeducational school in an outer suburb of Melbourne participated in the study.

Materials and procedure

A pool of algebraic story problems that had been found in previous studies of Australian tenth grade students to have pass rates between 40 and 60 per cent was assembled. This pool was drawn upon in developing the five exercises that comprised the test battery. The exercises differed in task requirements as well as in format, and they were presented to all students in the order listed below.

Exercise A contained eight problems, each of which lacked a component of information that was essential if the problem was to be solved. An open-ended format was employed, and the instructions required the student to write on the

answer booklet the additional information that the problem had to contain in order for solution to be possible. Two examples are:

A1: "Amanda is older than Brenda. If their combined ages amount to 28 years, how old is Brenda?"

A2: "If the size of one angle of a triangle is twice the size of a second angle (A), what is the size of A (in degrees)?"

The eight problems in Exercise B were similar in terms of problem category and specific detail to the problems in Exercise A, but in this case each problem contained not only information that was necessary for solution but also information that was irrelevant for solution. The student's task was to underline on the test booklet the information within the problem (which always was a single sentence or phrase) that was not needed in order to solve the problem. Two examples are:

B1: "Amanda is four years older than Brenda. Brenda is two years older than David. If the combined ages of Amanda and Brenda amount to 28 years, how old is Amanda?"

B2: "If the size of one angle of a triangle is twice the size of a second angle (A), and the size of the third angle is three times the size of the second angle, and the right angle of a triangle is 90 degrees, what is the size of angle A (in degrees)?"

In Exercise C the students received the same eight problems as in Exercise B, including the irrelevant information, but the requirement in this case was that each problem be solved. Multiple-choice format was used. The four response options available for each question included the correct answer, a solution that probably would have been reached through computational error, a solution that might have been reached through reliance on incomplete information, and a solution that could have been reflected use of the irrelevant information. An example of the test format in Exercise C is:

C1: "Amanda is four years older than Brenda. Brenda is two years older than David. If the combined ages of Amanda and Brenda amount to 28 years, how old is Amanda?"

- (a) 14 years old
- (b) 18 years old
- (c) 12 years old
- (d) 16 years old

Exercise D was designed to provide a measure of recognition memory. The students were presented with 16 problems, consisting of eight problems from Exercises A and B and eight further problems that had not been given beforehand. The task of each person was to make judgments on a five-point scale of certainty whether a problem was "old" (had been presented in Exercise A or) or "new" (was being presented now for the first time). Each "new" problem was written so that it matched an "old" problem in surface detail, but was dissimilar in problem structure (and hence in schematic knowledge requirements). "New" and "old" problems were presented in Exercise D in a randomly ordered sequence. In Exercise D A1 was an "old" problem, and the corresponding "new" problem was:

D1: "Amanda is four times as old as Brenda and Brenda is twice as old as David. What is the ratio of their ages?"

Similarly, A2 was presented as an "old" problem, and the corresponding "new" problem was:

D2: "The size of one angle of a right-angle triangle is 50 degrees. What is the size (in degrees) of the other angle?"

In Exercise E students were presented with five sets, each containing four algebraic story problems. Three of the problems required the same or comparable operations in order for the correct solution to be reached. Students were told to identify the problem within the set that they considered was different from the other three, and the instructions gave no indication as to what was meant by different. An example of this type of problem is:

- E1: (a) "A man drives along a road which is at an angle of 5 degrees to the horizon. He covers a distance of 20 km in 20 minutes. What is the speed of the car?"
(b) "How long must a ladder be so that it just rests against a vertical wall of height 2.5 m at an angle of 40 degrees to the horizontal?"
(c) "A crane of length 15 m can work to a maximum of 82 degrees to the horizontal. What is the maximum height to which the crane can lift a weight?"
(d) "A child's slide is 3 m long and has a vertical ladder 2.2 m long. Find the angle which the slide makes with the vertical."

Testing was in two sessions, which were separated by one day. In the first session all students completed Exercise A, then Exercise B, and finally Exercise C. In the second session Exercise D was completed prior to Exercise E. Instructions for each exercise were given verbally as well as in writing. Prior to commencing Exercises A, B, C, and E the students were given a sample problem, together with the correct answer. The timetabling of classes within the school made it necessary to impose a time limit of 15 minutes for the completion of each exercise. However, all but four of the 76 students finished each exercise within this time limit. The results from these four students are not included in the data analysis.

RESULTS

The mean success rates for students across the problems in the different exercises were 65 per cent for Exercise A (identification of missing information), 61 per cent for Exercise B (deletion of irrelevant information), 56 per cent for Exercise C (problem solution), 76 per cent for Exercise D (recognition memory), and 55 per cent for Exercise E (categorisation as indexed through oddity discrimination). Table 1 shows the intercorrelations between the total scores (number correct) of students on the five exercises. There was a correlation of $+0.89$ between the ability of students to identify missing information and to delete irrelevant information in algebraic story problems, and there was an even higher correlation between the skill with which students could delineate the necessary and sufficient information for problems to be soluble and the likelihood that the students could provide the correct answer to these problems. Problem categorisation as identified through oddity discrimination also correlated highly with solution rates and the two measures of problem editing, but scores on recognition memory did not significantly with any measure other than problem categorisation.

TABLE 1
PRODUCT-MOMENT CORRELATIONS BETWEEN SCORES ON THE FIVE EXERCISES

	A	B	C	D	E
Exercise A (missing information)	—	0.89*	0.91*	0.24	0.68*
Exercise B (irrelevant information)		—	0.94*	0.27	0.72*
Exercise C (problem solution)			—	0.29	0.75*
Exercise D (recognition memory)				—	0.43*
Exercise E (oddy discrimination)					—

* $P < 0.01$

The relationship between the five performance measures were studied further through stepwise multiple regression analysis. The scores of students on Exercise C (problem solution) were taken as the criterion measure, and the predictor variables were scores on Exercises A, B, D, and E as well as the sex of students. The number of problems in Exercise B on which students had been successful in detecting irrelevant information accounted for 87.6 per cent of variance in the number of problems in Exercise C that the students had solved accurately. The ability to specify missing information (Exercise A) and to detect the odd problem within a set of problems (Exercise E) accounted for additional variance of 2.8 per cent and 0.7 per cent respectively. Neither the recognition memory scores nor the sex of the students accounted for significant variance in problem solution rate over and above that attributable to identification of missing information, deletion of irrelevant information, and problem categorisation.

TABLE 2
SUMMARY OF MULTIPLE REGRESSION ANALYSIS WITH EXERCISE C
(SOLUTION) AS THE CRITERION MEASURE

Step	Predictor	R^2	ΔR^2	β	F
1	Exercise B (irrelevant information)	0.88	—	0.94	501.43**
2	Exercise A (missing information)	0.90	0.02	0.37	21.51**
3	Exercise E (oddy discrimination)	0.91	0.01	0.13	6.37*

* $P < 0.01$, ** $P < 0.001$

The interrelationships that have been identified through analysis of total scores (summed over problems within an exercise) are also evident when consideration is limited to performance on a specific problem. It was noted earlier that the same eight problems, but in different format and with different task requirements, were employed in Exercises A, B, C and D (as an example, see problems A1, B1, C1, and D1 in the Method section). Since the response a student gave to each problem within each exercise was scored as correct or incorrect, it is possible to establish on a problem-by-problem basis whether identification of missing information and deletion of irrelevant information in problems are necessary and sufficient skills for solution of these same problems and for recognition memory.

The cross-tabulations in Table 3 show for each of the eight problems the number of students in the sample demonstrating specific patterns of performance across exercises. Despite Exercise C having been in multiple choice format (with the possibility that a correct response could be achieved simply through guessing), it is clear from Table 3 that for each problem the students who gave the correct solution were almost always those who had given a correct response on the same problem in Exercises A and B. On problem 1, for example, 37 of the 72 students reached the correct solution, and 33 of these 37 were also correct on the corresponding problem in Exercises A and B. The 35 students who reached the wrong solution on problem 1 included only one subject who had been correct on the corresponding problem in Exercises A and B. More importantly, not a single person who failed on both the detection of missing information and the identification of irrelevant material solved the problem correctly. Similar relationships were obtained for all eight problems. Overall, students who gave the correct response to a problem in Exercises A and B had a 95.6 per cent likelihood of providing the correct solution to this same problem in Exercise C. The success likelihood for solution was 21.2 per cent if an error had been made in one of Exercises A and B, and only 0.6 per cent if an innocent response to the problem had been given in *both* Exercises A and B.

TABLE 3

PERFORMANCE ACROSS PROBLEMS ON EXERCISES C (SOLUTION) AND D (RECOGNITION MEMORY) AS A FUNCTION OF PERFORMANCE ON EXERCISES A (IDENTIFICATION OF MISSING INFORMATION) AND B (DELETION OF IRRELEVANT INFORMATION)

Performance A B		Problem																Total
		1		2		3		4		5		6		7		8		
		C+ C-	C+ C-	C+ C-	C+ C-	C+ C-	C+ C-	C+ C-	C+ C-	C+ C-	C+ C-	C+ C-	C+ C-	C+ C-	C+ C-	C+ C-	C+ C-	
+	+	33	1	52	1	28	1	52	6	39	1	46	0	14	1	38	3	302 14
+	-	2	4	0	10	2	23	0	2	0	4	2	2	1	2	0	2	7 49
-	+	2	1	2	1	0	3	1	4	3	5	0	5	2	0	2	3	12 22
-	-	0	29	0	6	1	14	0	7	0	20	0	17	0	52	0	24	1 169
		1	2	3	4	5	6	7	8	Total								
		D+ D-	D+ D-	D+ D-	D+ D-	D+ D-	D+ D-	D+ D-	D+ D-	D+ D-	D+ D-	D+ D-	D+ D-	D+ D-	D+ D-	D+ D-	D+ D-	
+	+	30	4	38	15	25	4	42	16	30	10	41	5	10	5	40	1	256 60
+	-	6	0	5	5	19	6	1	1	2	2	3	1	2	1	2	0	40 16
-	+	1	2	1	2	2	1	4	1	7	1	4	1	1	1	4	1	24 10
-	-	23	6	5	1	10	5	4	3	15	5	13	4	30	22	18	6	118 52

+ signifies correct response, and - incorrect response

In contrast to the close relationship found between text editing and problem solution, success at recognition memory was nowhere near as dependent on the ability to undertake text editing. Table 3 shows interrelations in performance on Exercises A, B, and D for each of the eight problems. It is clear that students who were unable to identify missing information *or* delete irrelevant information in a problem had been almost as readily able to encode the structure of the problem (as indexed by their ability to distinguish this "old" problem from its "new" counterpart under the recognition memory conditions employed in Exercise D) as the students who had given the correct response to the problem on Exercises A and B.

The likelihood that students who failed on both Exercises A and B would succeed on recognition memory was 69.5 per cent, and the comparable value for those who had succeeded on both Exercises A and B was 81.0 per cent. Thus on three-quarters of times that students were unsuccessful in text editing, they nevertheless remembered enough about the problem to be able on the next day to distinguish this problem from another problem that had similar surface content but different structure. Not only was there adequate memory for problems in the absence of text editing but in approximately one in five occasions students failed on memory for problems on which they had edited the text successfully.

The relationship between problem solution rates and recognition memory scores is not reported in Table 3. However, in keeping with the low product-moment correlation reported in Table 1 for total scores on these two variables, the median phi coefficient for success on problem solution and success on recognition memory was 0.09 across the eight problems. In contrast, the median phi coefficient was 0.79 for Exercises A and C, and 0.84 for Exercises B and C.

It was noted earlier that the multiple choice options provided in Exercise C, where the task required solution of problems, were constructed with the objective of identifying the types of errors that students made. Across all eight problems there were 254 errors in solution. Of these, 201 involved choice of the response option that incorporated the irrelevant information included in the problem, 37 were classified as mistaken computation, and 16 entailed use of only part of the information that was available within the problem.

DISCUSSION

The ability of students to edit the text of algebraic story problems proved to be highly predictive of problem solution rates. Across the eight problems and 72 students, there was only one instance of a student being correct in solution after having failed both to identify what essential information was missing and to delete the information that was irrelevant when the same problem was presented in these different test formats. In contrast, in 302 of the 322 occasions where there was correct solution of a problem students had been able to edit the text for this problem both to identify missing information and to delete irrelevant information.

Problem sorting and categorisation have been used as an index of schematic knowledge in earlier research (for example, Hinsley *et al.*, 1977; Silver, 1979). Categorisation was assessed in the present study by requiring students to identify which problem in a set of four problems was different from the other three. Consistent with the result reported by Silver (1979), the ability to solve algebraic text problems correlated with the ability to categorise similar problems. Of more specific interest, those students who could determine whether algebraic story problems contained the information that was necessary and sufficient for solution could in general not only solve these same problems but determine whether problems were similar to, or different from, other problems in terms of the mathematical principles that were involved. The correlations reported in Table 1 indicate that students who lacked success in categorising problems also performed poorly in editing and in solving problems.

Whereas there was a close relationship between problem categorisation, editing, and solution, memory for problems did not correlate significantly with either editing or solution scores, and the correlation between memory and problem categorisation was only 0.43. In studies that have taken memory for algebraic story problem as an index of schematic knowledge (for example, Silver, 1981;

Mayer, 1982), memory has been measured through recall rather than recognition. In the present study, however, recognition was assessed rather than recall, primarily because group testing together with time limitations on access to students made a recognition test easier to administer.

Recognition memory was assessed by requiring students to distinguish between problems which had been presented earlier ("old" problems) and those now being presented for the first time ("new" problems). Pairs of "old" and "new" problems were written with the intention that they would be similar in surface detail, but dissimilar in problem structure. However, recognition memory scores did not correlate significantly with either problem editing or problem solution scores. The students were accurate in recognition memory for 80·5 per cent of problems for which they gave the correct solution, but they were also accurate in recognition memory for 70·5 per cent of problems for which they had provided an incorrect solution. It may be that the "old" and "new" problems used in assessing recognition memory differed with respect to problem detail, and not just problem structure. Silver (1981) found that although students who had performed poorly on problem solution generally could not recall problem structure, they were often able to remember contextual information and the question posed in the problem. Relationships between memory, problem editing, and problem solution should probably be assessed by measuring memory in terms of recall rather than recognition.

The five exercises were completed in the same order by all students. Although oddity discrimination could have been scheduled at any stage in testing, task requirements imposed some constraints on sequencing of the other exercises. For example, memory could be assessed only after presentation of the material to be remembered. This material was provided in the solution exercise. Since the text editing and solution exercises involved the same algebraic problems, text editing was completed prior to solution to reduce the likelihood of contamination of processes. In support of the claim that text editing is a prerequisite for solution rather than vice-versa, students were often successful on text editing and not on solution but they rarely were successful on solution and not on text editing. At the same time it would be interesting to establish whether testing for solution prior to, rather than after, text editing facilitates performance on text editing.

Since the eight problems in Exercises A, B, and C involved different schematic knowledge, it is not possible to establish from the composite scores used in data analysis whether text editing is a specific or a general skill. Mayer (1981) has emphasised the domain-specific nature of schematic knowledge associated with algebraic story problems. Through analysis of textbook content, he identified 25 general families of problems (for example, motion, current, interest, rectangle, percent, probability), and within each family a number of different formats or templates (such as, in the case of motion problems, overtake, closure, speed change, and round trip). To determine the extent to which text editing is a general skill, it would be necessary to contrast performance within and between formats and templates as well as within and between general families of problems. A question of further interest is whether text editing is a context-specific skill or a more general capability. Perhaps the students in the present study who were most successful in editing algebraic story problems would also have excelled in identifying missing and irrelevant information in non-mathematical text. A general conceptual ability may be involved.

Some students succeeded on text editing but failed on problem solution, but it was rare for a student to fail on text editing but succeed on problem solution. Such a result is consistent with the proposition that successful text editing is a

precursor to problem solution, with errors in solution reflecting mistakes in calculation and other post-editing processes. The data, however, simply indicate that editing and solution are correlated, and not that editing is a pre-solution process. When attempting Exercise A students may not have concluded that there were insufficient givens simply from analysis of the text and through knowledge of problem templates. Students may come to realise that essential information is missing only when they cannot reach a solution on the basis of what is available in the text. Similarly, when presented with Exercise B students may attempt solution and then detect that specific information is irrelevant on the grounds that part of the text does not have to be used in achieving a solution to the problem. The issue of whether text editing and problem solution occur serially or in parallel can be studied by requiring subjects to provide verbal accounts of their method of approach to tasks such as Exercises A, B, and C. Since the basis for editing the text of algebraic story problems may vary with level of mathematical competence, novices and experts should be compared.

Finally, in the conventional teaching of mathematics greater emphasis has been placed on practice and rehearsal of explicit solution processes than on the development and acquisition of schematic knowledge. It has been demonstrated in the present study that competence in problem solution is highly correlated with competence in text editing. The nature of the processes underlying this correlation can perhaps be elucidated by determining whether a direct focus on the acquisition of text editing skills during instruction facilitates problem solution.

ACKNOWLEDGMENTS.—The authors are grateful to the Australian Council for Educational Research and to P. Pattison of the University of Melbourne for access to the item banks from which most of the algebraic problems used in the research were derived. Thanks are also due to students and teachers at the school where the testing was undertaken.

Correspondence and requests for reprints should be sent to Professor Ray Over, Department of Psychology, La Trobe University, Bundoora, Australia 3083.

REFERENCES

- CHI, M. T., FELTOVICH, P. J. H., and GLASER, R. (1981). Categorization and representation of physics problems by experts and novices. *Cogn. Sci.*, 5, 121-152.
- DE JONG, T., and FERGUSON-HESSLER, M. G. M. (1986). Cognitive structures of good and poor novice problem solvers in physics. *J. educ. Psychol.*, 78, 279-288.
- HINSLEY, D. A., HAYES, J. R., and SIMON, H. A. (1977). From words to equations: meaning and representation in algebra word problems. In JUST, M. A. and CARPENTER, P. A. (Eds.), *Cognitive Processes in Comprehension*. Hillsdale, NJ: Lawrence Erlbaum.
- LARKIN, J., McDERMOTT, J., SIMON, D. P., and SIMON, H. A. (1980). Expert and novice performance in solving physics problems. *Science*, 208, 1335-1342.
- MAYER, R. E. (1981). Frequency norms and structural analysis of algebra story problems into families, categories, and templates. *Instruct. Sci.*, 10, 135-175.
- MAYER, R. E. (1982). Memory for algebra story problems. *J. educ. Psychol.*, 74, 199-216.
- MAYER, R. E. (1983). *Thinking, Problem Solving, Cognition*. San Francisco; W. H. Freeman.
- SCHOENFELD, A. H., and HERMANN, D. J. (1982). Problem perception and knowledge structure in expert and novice mathematical problem solvers. *J. exp. Psychol.: Learn., Memory, Cogn.*, 8, 484-494.
- SILVER, E. A. (1979). Student perceptions of relatedness among mathematical verbal problems. *J. Res. Math. Educ.*, 10, 195-210.
- SILVER, E. A. (1981). Recall of mathematical problem information: solving related problems. *J. Res. Math. Educ.*, 12, 54-64.

(Manuscript received 29th September, 1987)

GENDER-STEREOTYPIC PERCEPTIONS OF ACADEMIC DISCIPLINES

By JOHN ARCHER AND SARA FREEDMAN
(*School of Psychology, Lancashire Polytechnic, Preston*)

SUMMARY. Sixty college students, aged 16-20 years, rated 10 academic disciplines along seven 7-point dimensions including masculine-feminine, in a $2 \text{ (sex)} \times 3 \text{ (academic background)}$ factorial design. Engineering, physical sciences, and mathematics were rated as significantly masculine whereas English, biology, psychology, French and sociology were rated as significantly feminine. There was no effect of the sex of the rater, and only in the case of biology was there a significant influence of academic background. For the mean ratings for each discipline, masculine-feminine was correlated with several other dimensions, but a stepwise regression revealed only difficult-easy as a significant predictor. On the individual ratings for three specific disciplines, correlations between masculine-feminine and other dimensions were small and non-significant, and the regression analyses were all non-significant. These results cast doubt on a previous conclusion that there is a *cluster* of other attributes associated with "masculine" when academic disciplines are rated.

INTRODUCTION

THERE has been considerable interest in the role of pupils' stereotypic attitudes in contributing to female underachievement in science and mathematics (e.g., Blackstone and Weinreich-Haste, 1980; Kelly, 1981a, 1981b). There are, however, few systematic investigations of the gender-stereotypic connotations of scholastic disciplines. Ormerod (1981) studied preferences for different disciplines among a nationwide sample of 14-year-old British pupils. The "gender" of each discipline was determined by calculating the difference between the sexes' preference for that subject: chemistry, physics and mathematics were chosen most frequently by boys, and religious education, English and French were chosen most frequently by the girls.

A more direct method of assessing the gender stereotypic connotations of scholastic disciplines was used by Weinreich-Haste (1979). British undergraduates rated academic and practical disciplines on a 6-point semantic differential with seven dimensions including masculine-feminine. The means for engineering, physics and mathematics were nearest to the masculine pole and modern languages, cookery and sociology were nearest the feminine pole. No differences were found between male and female students, or between those from social science and other academic backgrounds. Intercorrelations between the masculine-feminine ratings and other dimensions, such as science-arts and hard-soft, were interpreted as indicating that masculine has connotations of hard, complex, intellect-based and scientific, when applied to scholastic disciplines. Representation of the mean ratings for masculine-feminine and science-arts as orthogonal dimensions was interpreted as evidence for a "cluster" of masculine scientific subjects.

In a similar study (Weinreich-Haste, 1981), 13-14-year-old children rated 10 school disciplines along 10 dimensions. Woodwork received the most masculine mean rating, followed by physics and chemistry; cookery received the most feminine rating, followed by typing, English and French. On the basis of significant

intercorrelations between ratings along the different dimensions, the conclusion was again drawn that there was a cluster of attributes applied to science, including masculine, hard and complex.

Although providing interesting data on the gender-stereotypic connotations of scholastic disciplines, there are a number of methodological problems with these studies. The first is the lack of a consistent criterion for deciding whether each discipline is masculine, feminine or neutral. Such a decision can, however, be readily reached by incorporating a neutral point on the scale, and using a t-test to determine whether the ratings are significantly different from the neutral or midpoint (Walker *et al.*, 1986).

The second problem concerns the assessment of the relationships between the different dimensions. In one report (Weinreich-Haste, 1979) intercorrelations between the dimensions were much higher than in the other one (Weinreich-Haste, 1981). It is apparent that in the first study the *mean* ratings for each academic discipline were used in calculating correlations between the dimensions, whereas in the second study individual values were used. The use of individual values poses a statistical problem: either the individual respondents' ratings for each academic discipline are considered separately or else the 10 sets of correlations for each respondent (one for each discipline) are treated as independent data points. Since no separate correlation matrices were presented for different disciplines, it seems likely that Weinreich-Haste (1981) used the second method, which confounds within and between subject variations.

The present paper reports an investigation of the gender-stereotypic connotations of scholastic disciplines which seeks to overcome these methodological limitations, and to investigate systematically the influence of the rater's sex and academic background on the judgments. Respondents rated the disciplines along 7-point dimensions, and whether the ratings for each discipline were different from the midpoint was tested statistically. Male and female respondents from three academic backgrounds were used. Intercorrelations were calculated, first between the mean values for each academic discipline, and secondly between the individual ratings on the dimensions for three disciplines, one masculine, one feminine and one neutral. Multiple regressions were also computed to explore further the relations between masculine-feminine and other dimensions.

METHOD

Respondents

The respondents were 30 male and 30 female A-level college students (ages 16-20 years) from tertiary colleges in Bolton and Bury, Greater Manchester. A third of the sample was taking foreign modern languages, a third the physical sciences and a third psychology.

Procedure

Each student was asked to rate 10 academic disciplines with respect to the following dimensions: difficult-easy, interesting-boring, useless-useful, masculine-feminine, simple-complicated, and science-arts. Each dimension consisted of seven points, the midpoint being neither one nor the other, and the three on each side being slightly, quite and extremely.

Design and data analysis

For each academic discipline, all 60 scores were tested for a significant departure from the neutral or midpoint on the masculine-feminine dimension, using a

one-sample t-test. Then a 2 (sex) $\times$ 3 (academic background) ANOVA (10 respondents in each cell) was carried out on the masculine-feminine dimension for each academic discipline. Thirdly, the individual masculine-feminine ratings were intercorrelated, and regressed on the five other dimensions separately for three different disciplines, one masculine, one feminine and the other neutral. Finally, the *mean* masculine-feminine ratings for each academic discipline were intercorrelated and regressed on the mean ratings on the other dimensions.

RESULTS

Gender connotations

Table 1 shows the mean ratings for all 60 students for the various academic disciplines along the masculine-feminine dimension, together with whether they differ from neutral. The first four disciplines are significantly different from neutral in the masculine direction, and the last five in the feminine direction, German being the only one rated as neutral.

TABLE 1
MEAN MASCULINE-FEMININE RATINGS OF ACADEMIC DISCIPLINES FOR ALL 60 STUDENTS

Academic discipline	Mean rating	Difference from Neutral (t-value)	Significance level P
Engineering	2.15	-11.76	<0.0001
Physics	2.73	-8.24	<0.0001
Chemistry	3.23	-6.67	<0.0001
Maths	3.52	-4.29	<0.0001
German	4.17	1.49	NS
English	4.32	3.38	<0.001
Biology	4.32	2.75	<0.001
Psychology	4.40	4.06	<0.0001
French	4.42	4.10	<0.0001
Sociology	4.43	3.31	<0.001

TABLE 2
MEAN SCIENCE-ARTS RATINGS OF ACADEMIC DISCIPLINES FOR ALL 60 STUDENTS

Academic discipline	Mean rating	Difference from Neutral (t-value)	Significance level P
Physics	1.12	-59.97	<0.0001
Chemistry	1.18	-38.46	<0.0001
Biology	1.55	-20.05	<0.0001
Engineering	2.10	-14.20	<0.0001
Maths	2.38	-10.61	<0.0001
Psychology	3.15	-5.34	<0.0001
Sociology	3.77	-1.46	NS
French	5.43	9.86	<0.0001
German	5.55	9.36	<0.0001
English	5.73	11.55	<0.0001

For comparative purposes, Table 2 shows a similar analysis carried out on the science-arts dimension. In this case, only sociology was rated as neutral, whereas the top six disciplines were rated as significantly different in the science direction and the last three as significantly different in the arts direction.

Analysis of variance

The 2×3 ANOVA carried out on the masculine-feminine scores for each of the 10 dimensions showed no significant main effects of sex of respondent, and only in the case of biology was there a main effect of academic background ($F[2,54] = 8.09$ $P < 0.001$). Scheffé tests revealed a highly significant difference between language and physical science students ($F[2,54] = 16.3$; $P < 0.001$), and a lesser difference between language and psychology students ($F[2,54] = 4.88$; $P < 0.05$) and between physical science and psychology students ($F[2,54] = 3.27$; $P < 0.05$). Further analysis showed that the language students did not rate biology as significantly different from neutral (mean = 3.8) whereas both the psychology and physical science students did (means, 4.35 and 4.8 respectively; $P < 0.01$ and < 0.001).

There were also sex $\times$ academic background interactions for the ratings of chemistry ($F[2,54] = 3.48$; $P < 0.05$) and engineering ($F[2,54] = 7.15$; $P < 0.01$). In view of the increased possibility of Type 1 errors with the repeated ANOVAs carried out in the present study, only the latter, more highly significant, interaction is described. It resulted from engineering being rated as significantly more masculine by female than male psychology students and significantly more masculine by male than female science students. For language students, there was no difference between the ratings of male and female students.

Intercorrelations of mean scores

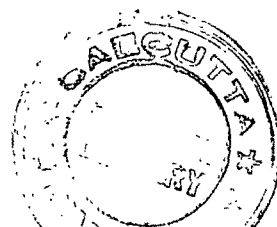
Table 3 shows the intercorrelations between the six dimensions for the mean scores from the different disciplines. In this case, masculine-feminine showed significant correlations with difficult-easy (i.e., masculine was related to difficult), interesting-boring (masculine related to boring) and simple-complex (masculine related to complex).

TABLE 3

INTERCORRELATIONS BETWEEN THE SIX DIMENSIONS FOR MEAN RATINGS ON THE 10 DISCIPLINES

	Masculine-feminine	SA	UU	DE	IB
Science-arts (SA)	0.624*				
Useless-useful (UU)	-0.477	-0.413			
Difficult-easy (DE)	0.734**	0.759**	-0.204		
Interesting-boring (IB)	-0.611*	-0.196	0.183***	-0.513	
Simple-complex	-0.668*	-0.653*	0.203	-0.967***	0.562

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$



Intercorrelations of individual scores

Tables 4, 5 and 6 show the intercorrelations between the six dimensions carried out on the individual scores, for engineering, sociology and German respectively. In each case, none of the correlations between masculine-feminine and other dimensions is statistically significant.

TABLE 4
INTERCORRELATIONS BETWEEN RATINGS ON THE SIX DIMENSIONS FOR ENGINEERING

	Masculine-feminine	SA	UU	DE	IB
Science-arts (SA)	0.153				
Useless-useful (UU)	-0.055	-0.240			
Difficult-easy (DE)	-0.031	0.382**	-0.119		
Interesting-boring (IB)	0.103	-0.085	-0.203	-0.196	
Simple-complex	-0.074	-0.348**	0.287*	-0.559***	0.058

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$

TABLE 5
INTERCORRELATIONS BETWEEN RATINGS ON THE SIX DIMENSIONS FOR SOCIOLOGY

	Masculine-feminine	SA	UU	DE	IB
Science-arts (SA)	-0.215				
Useless-useful (UU)	-0.247	-0.239			
Difficult-easy (DE)	0.025	0.223	-0.205		
Interesting-boring (IB)	-0.245	0.176	-0.508***	0.116	
Simple-complex	-0.006	-2.210	0.291*	-0.673***	-0.232

* $P < 0.05$; *** $P < 0.001$

Regression analyses

Regression of the mean values on masculine-feminine for the 10 academic disciplines on those for the other five dimensions produced a value of 78.0 per cent for R^2 (adjusted $R^2 = 50.5$ per cent). However, despite this high value for R^2 , the F value was not significant, in view of the low numbers of degrees of freedom in the error term compared with those in the predictors ($F = 2.838$; 5, 4df, NS). A stepwise regression on the same data did reveal a significant entry value for difficult-easy ($R^2 = 53.85$) in step 1 ($F = 9.36$), but no further values reached criterion.

TABLE 6
INTERCORRELATIONS BETWEEN RATINGS ON THE SIX DIMENSIONS FOR GERMAN

	Masculine-feminine	SA	UU	DE	IB
Science-arts (SA)	0.220				
Useless-useful (UU)	0.086	0.065			
Difficult-easy (DE)	0.019	0.024	0.214		
Interesting-boring (IB)	0.182	0.227	-0.500***	-0.376**	
Simple-complex	0.005	0.141	-0.049	-0.388**	0.260*

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$

The results of the multiple regressions carried out on the individual scores were as follows. For sociology, regressing masculine-feminine scores on the other five dimensions produced an R^2 of 17.2 per cent (adjusted $R^2 = 9.5$ per cent), which approached but did not reach significance ($F = 2.236$; 5,54df; NS). For engineering, $R^2 = 4.9$ per cent (adjusted value 0 per cent), which was non-significant ($F = 0.055$; 5,54df). For German, $R^2 = 10.3$ per cent (adjusted = 2.0 per cent), which was again non-significant ($F = 1.24$; 5,54df).

DISCUSSION

With the exception of German, all the academic disciplines used in the present study were significantly different from neutral, i.e., were stereotyped as either masculine or feminine. In agreement with Weinreich-Haste (1979, 1981), engineering, physics, chemistry and maths were all viewed as masculine. French and sociology were the most feminine, English, biology and psychology also being viewed as feminine, again in broad agreement with Weinreich-Haste (1979). These results are also consistent with the ratio of boys' to girls' pass rates in the disciplines (e.g. Murphy, 1979; CSO, 1985), and the numbers of each sex choosing them at higher academic levels (e.g., Ormerod, 1981).

Although there was a correspondence between disciplines rated as masculine and as scientific for the physical sciences, maths and engineering, two other disciplines — biology and psychology — were viewed as feminine and scientific. Kelly and Smail (1986) found that there was a small but significant correlation between a feminine self-image and an interest in natural history for both sexes in a large sample of British 11-year-olds. Other data from the same sample (Smail and Kelly, 1984) showed that girls were more interested in natural history and human biology whereas the boys were more interested in the physical sciences.

The gender ratings obtained in the present study were similar for male and female raters, which is consistent with Weinreich-Haste's (1979) student sample, but not with her data from schoolchildren (Weinreich-Haste, 1981), where boys tended to rate most subjects more towards the masculine pole. The analysis of

variance in the present study found the following influence of academic background: language students perceived biology as a neutral rather than a feminine subject, in contrast to psychology and physical science students.

The intercorrelations between the mean scores showed that masculine-feminine was significantly positively correlated with science-arts and difficult-easy, and it was negatively correlated with interesting-boring and simple-complex. Although different dimensions were used by Weinreich-Haste (1979), those for science-arts and simple-complex were similar in the two studies, and the overall pattern of high correlations between masculine-feminine and other dimensions for the mean scores was replicated. However, the stepwise regression showed that only difficult-easy was a significant predictor for masculine-feminine. Thus the earlier conclusion that a *cluster* of attributes is associated with masculine-feminine (Weinreich-Haste, 1979) was likely to have been premature.

The intercorrelations between individual scores on the three disciplines chosen to represent masculine, feminine and neutral, showed no significant correlations between masculine-feminine and any other dimension; the scores on other dimensions did not significantly predict those on masculine-feminine in the regression analyses. These findings show that, once the variation between the academic disciplines has been removed, masculine-feminine ratings are not related to those on the other dimensions.

The present results suggest, therefore, that the pattern of relationships between masculine-feminine and other dimensions found in Weinreich-Haste's (1981) study (where between- and within-subject variation were confounded) was produced because judgments vary in a predictable way with the target discipline, but not with individual raters for any specific discipline. Thus an individual who views physics as more masculine than another person does will not necessarily also view it as more scientific and more difficult; but if a particular academic discipline is rated as more masculine than another discipline, it will also be viewed as more scientific and more difficult. However, even this conclusion requires qualification since the multiple regression analysis in the present study indicated that only the difficult-easy dimension was a significant predictor of masculine-feminine. This narrows down the implicit connotations of masculine-feminine, so that future studies could concentrate on exploring this particular link.

Two additional points concern the sample and methodology used in the present study. The sample size was smaller than in the previous studies, and restricted to A-level college students. However, since comparable results were obtained, and the sample size was sufficient for statistical analysis, this is unlikely to have substantially affected the findings.

Regarding the methodology, although the semantic differential provides a systematic way of investigating the connotations of particular target words it does constrain the possible choices. First, the use of bipolar dimensions obscures the possibility that some disciplines may have *both* masculine and feminine connotations (cf. Bem, 1974; Spence *et al.*, 1975). A second, and more serious, constraint is the prior assumption that the specific dimensions are necessarily relevant to the target words. A third is the assumption that the cognitive structure of consensual beliefs about the target word is adequately represented by a single dimension. Fourthly, rating scales provide a static view of stereotypes, neglecting their use in social interactions.

Rating scales involving separate masculine and feminine dimensions would resolve the first issue. The second constraint can be overcome by using repertory grids (e.g., Baldwin *et al.*, 1986) or introspective methods (Zavalloni, 1971), where

personally-relevant adjectives are derived from the individual's responses. The third one requires a method whereby respondents judge the probability that descriptive adjectives are associated with the target person (Deaux and Lewis, 1983). The final constraint can be overcome by investigating how stereotypic judgments are used in conversations (Condor, 1987; Wetherell *et al.*, 1987). However, no one method by itself is likely to lead to a full understanding of gender stereotyping.

Correspondence and requests for reprints should be addressed to John Archer, School of Psychology, Lancashire Polytechnic, Preston, Lancs., PR1 2TQ.

REFERENCES

- BALDWIN, A. C., *et al.*, (1986). Androgyny and sex role measurement: a personal construct approach. *J. Person soc. Psychol.*, 51 (5), 1081-1088.
- BEM, S. L. (1974). The measurement of psychological androgyny. *J. consult. clin. Psychol.*, 42 (2), 155-162.
- BLACKSTONE, T., and WEINREICH-HASTE, H. (1980). Why are there so few women scientists and engineers? *New Society*, 51, 383-385.
- CENTRAL STATISTICS OFFICE (1985). *Social Trends 1984*. London: HMSO.
- CONDOR, S. (1987). From sex categories to gender boundaries: reconsidering sex as a "stimulus variable" in social psychological research. *British Psychological Society Social Psychology Section Newsletter (Special Issue: Women and Psychology)*, 17, 48-71.
- DEAUX, K., and LEWIS, L. L. (1983). Assessment of gender stereotypes: methodology and components. *Psychol. Doc.*, 13, no. 25, cat. no. 2583.
- KELLY, A. (1981a). Sex differences in science achievement: some results and hypotheses. In KELLY, A. (Ed.), *The Missing Half: Girls and Science Education*. Manchester: Manchester University Press.
- KELLY, A. (1981b). Science achievement as an aspect of sex roles. In KELLY, A. (Ed.), *The Missing Half: Girls and Science Education*. Manchester: Manchester University Press.
- KELLY, A., and SMAIL, B. (1986). Sex stereotypes and attitudes to science among eleven-year-old children. *Br. J. educ. Psychol.*, 56, 158-168.
- MURPHY, R. J. L. (1979). Sex differences in examination performance: do these reflect differences in ability or sex-role stereotypes? In HARTNETT, O., BODEN, G., and FULLER, M. (Eds.), *Women: Sex Role Stereotyping*. London: Tavistock.
- ORMEROD, M. B. (1981). Factors differentially affecting the science subject preferences, choices and attitudes of girls and boys. In KELLY, A. (Ed.), *The Missing Half: Girls and Science Education*. Manchester: Manchester University Press.
- SMAIL, B., and KELLY, A. (1984). Sex differences in science and technology among 11-year-old schoolchildren: II-affective. *Res. Sci. technol. Educ.*, 2, 87-106.
- SPENCE, J. T., HELMREICH, R., and STAPP, J. (1975). Ratings of self and peers on sex role attributes and their relation to self-esteem and conceptions of masculinity and femininity. *J. Person soc. Psychol.*, 32 (1), 29-39.
- WALKER, P., SMITH, S., and LIVINGSTONE, A. (1986). Predicting the appropriateness of a typeface on the basis of its multimodal features. *Inform. Des. J.*, 5, 29-42.
- WEINREICH-HASTE, H. (1979). What sex is science? In HARTNETT, O., BODEN, G., and FULLER, M. (Eds.), *Women: Sex Role Stereotyping*. London: Tavistock.
- WEINREICH-HASTE, H. (1981). The image of science. In KELLY, A. (Ed.), *The Missing Half: Girls and Science Education*. Manchester University Press.
- WETHERELL, M., STIVEN, H., and POTTER, J. (1987). Unequal egalitarianism: a preliminary study of discourses concerning gender and employment opportunities. *Br. J. soc. Psychol.*, 59-71.
- ZAVALLONI, M. (1971). Cognitive processes and social identity through focussed introspection. *Eur. J. soc. Psychol.*, 1, 235-260.

(Manuscript received 27th October, 1988)

THE PREDICTIVE RELATIONSHIP BETWEEN BELIEFS, ATTITUDES, INTENTIONS AND SECONDARY SCHOOL MATHEMATICS LEARNING: A THEORY OF REASONED ACTION APPROACH

BY BRAHM NORWICH AND MARIANNE JAEGER
(*Institute of Education, University of London*)

SUMMARY. This study investigated how attitudes and intentions about learning mathematics might be related to subsequent mathematics learning and achievement using the Ajzen and Fishbein theory of reasoned action. The sample consisted of 142 boys and girls between 12 and 14 years old in a large inner city comprehensive school who were assessed in a follow-up design over a nine-month period. Beliefs about the outcomes of learning, attitudes to learning, perceptions of significant others' prescriptions about learning, intentions to engage in learning behaviours, self and teacher reported learning behaviour and mathematics achievement were assessed at both stages. Regression analysis suggested that while the expectancy-value components of attitude did relate to learning behaviour intentions, perceived prescriptions did not relate to intentions. There was a weak relationship between the two measures of learning behaviour, but with neither measure did intention independently predict future behaviour once prior behaviour was taken into account. The best predictor of subsequent mathematics achievement was prior achievement, though teacher-reported learning behaviour did have an independent relationship with subsequent achievement. The findings are discussed in terms of the assessment of learning behaviours, the relevance of the behaviour intention construct for repeated multiple behaviours and future work on how affective variables might be related to cognitive achievements.

INTRODUCTION

TEACHERS and researchers alike seem to stress the importance of pupils' attitudes, beliefs and interests in understanding school achievement. The relationship between attitudes and achievement is a popular area of study in research on science and mathematics education, yet this relationship appears to have modest empirical support. A meta-analysis of 123 studies examining the relationship between science attitudes and achievement in science (reported in Schibeci, 1984) found a mean correlation of only 0.11. In mathematics, researchers tend to find a positive, but moderate, relationship between attitudes and achievement, but that attitudes contribute little to the prediction of mathematics performance (Assessment of Performance Unit, 1981).

A number of criticisms have been suggested to account for the low to moderate correlations between attitudes and achievement. On methodological grounds, Schibeci (1984) points out that some researchers appear to take less care in measuring affective variables reliably and validly than in the measurement of the cognitive variables. He also notes that much of the research reports only simple bivariate relationships, ignoring the likelihood that more complex and subtle relationships between attitudes and achievement may exist. At a conceptual level, Schibeci suggests that much of the research lacks a theoretical framework concerning the nature of attitudes and the processes by which they relate to school learning and achievement. There is at best an explicit or implicit assumption that

attitudes to school or to school subjects should be related to achievement, if only on the grounds that positive attitudes lead to greater achievement. Of eight well-known theories of school learning reviewed by Haertel *et al.* (1983), several refer to affective factors, but only that of Bloom (1976) explicitly treats affective characteristics as both entry and outcome factors. Despite Bloom's emphasis on school and subject attitudes, little evidence beyond bivariate correlations between attitudes and achievement is presented.

Given the methodological and conceptual difficulties in this area of research, there is a need to consider theoretically *how* attitudes to subjects might be related to cognitive achievement in that subject. To do so, the assumption was made that participation in learning behaviours enables school learning and performance on measures of achievement, and it is proposed that these learning behaviours mediate between pupils' thoughts and feelings about school subjects and their achievement in that subject.

A theory of attitude-behaviour relationships, Ajzen and Fishbein's (1980) "theory of reasoned action" was elaborated in order to examine the relationship between learner attitudes and achievement. According to this model, the most immediate determinant of an individual's behaviour is his or her intention to perform that behaviour. Behavioural intention, in turn, is influenced by two factors, a prescriptive one and an evaluative one. The prescriptive factor, subjective norm, is the person's perception of the social pressure by important others to perform the behaviour or not. The evaluative factor, attitude to the behaviour, is the person's positive or negative affect towards the behaviour in question. The relative importance of attitudes and subjective norms may vary across behavioural domains and individuals, but it is assumed that they jointly determine behavioural intention.

The theory further permits examination of attitudes and subjective norms in terms of their underlying beliefs. Attitude towards behaviour is a function of the individual's beliefs about the outcomes of behaviour in question and his or her evaluations of these outcomes. Attitude can thus be estimated by weighting each outcome belief by its corresponding evaluation and then obtaining the sum of these products across all salient beliefs. Subjective norm is determined by normative beliefs concerning the likelihood that specific important others approve or disapprove of the behaviour for the individual and the individual's motivation to comply with these perceived prescriptions. Subjective norm is similarly estimated by weighting each normative belief by the corresponding motivation to comply, and obtaining the sum of these weighted beliefs across salient significant others.

There is a considerable amount of evidence for the theory in fields as diverse as consumer behaviours, voting and health related behaviours (cf. Ajzen and Fishbein, 1980; Manstead *et al.*, 1984). Other research, however, suggests that the posited relationships do not always obtain. Bentler and Speckart (1979) found that attitudes had a direct relationship with behaviour as well as an indirect one, through behavioural intention. Fredricks and Dossett (1983) found that attitude had a direct relationship with behaviour, not mediated by intention.

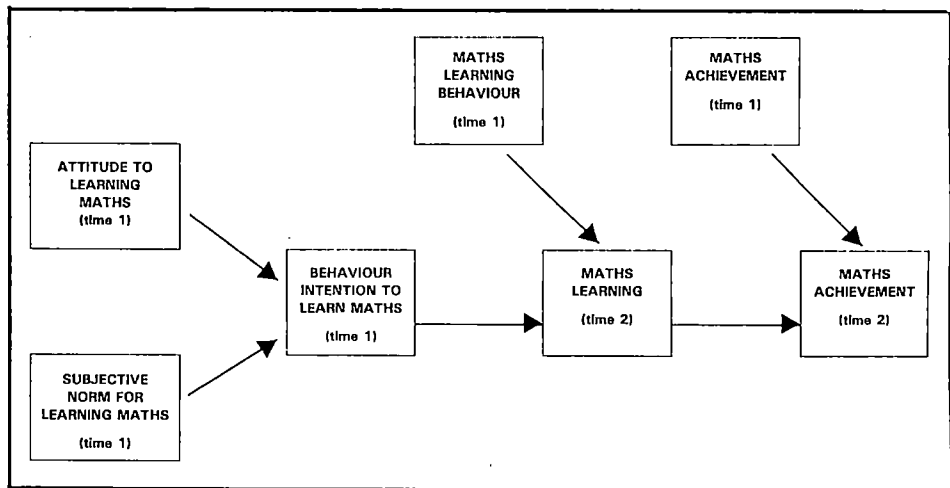
While recognising that there is a varying amount of empirical evidence for these posited relationships, it was nonetheless considered that the theory of reasoned action would provide a useful framework to define relevant concepts and to explore how affective factors might influence achievement. In elaborating the model to account for school achievement, account must be taken of the habitual and repeated nature of classroom learning behaviours, and that these behaviours and their outcomes involve skill and competence.

The theory of reasoned action represents an approach to the relationships between motivational and affective processes and school achievement which focuses on behavioural intention. It differs from other approaches to this field which either focus on the relationship between expectancy and achievement (Vollmer, 1986) or on attributions for success or failure and their effects on future behaviour (Butler, 1986; Clifford, 1986). In these approaches the central focus is either on the relationship between causal attributions for past performance and future expectancies or that between expectancies and future achievement. Saltzer (1982) has attempted to bring together the attributional and behavioural intention theories in a conceptual model. This is an elaboration of the Fishbein theory which incorporates attributions for the outcome of behaviours.

There is as yet little research evidence for this broader model. In constructing the model Saltzer has interpreted the influence of attributions for past outcomes as acting on outcome beliefs and evaluations. This is in line with the Fishbein analysis of the components of attitudes but is inconsistent with attributional theories which relate the stability dimension of causal attributions to expectancies of future behaviour (Weiner, 1986). The aim of the present study is to examine secondary school mathematics learning using only the theory of reasoned action adapted to the context of classroom learning (see Figure 1). In this model, mathematics achievement is determined by prior achievement and classroom learning behaviours in this subject. As the concept of learning behaviours is central to the present theoretical approach as a direct determinant of achievement and as a link to the young person's beliefs, attitudes and intentions, its role is explored in some detail in the present study.

FIGURE 1

MODEL OF RELATIONSHIPS BETWEEN AFFECTIVE FACTORS AND MATHEMATICS ACHIEVEMENT



METHOD

Participants

The participants were 66 boys and 76 girls from a large comprehensive school in an outer London borough. They were from two year groups: 74 second year

pupils, aged 12-13 and 68 third years, aged 13-14. The participants comprised about half of each year group and were selected from the first, third and fifth mathematics ability sets. The school was located in a socially disadvantaged area and drew its pupils from a multi-ethnic population.

Procedure

The research was carried out in two waves over a nine-month period in order to examine the various factors in the model with respect to possible changes in mathematics achievement over time. At time 1 and time 2 the attitudinal, normative and behavioural factors were assessed by means of questionnaire, and teachers' ratings provided an additional measure of pupils' learning behaviours. Two school based mathematics tests were used as measures of mathematics achievement, one for each year group. Just prior to the initial assessment of attitudes and achievement, the school had administered an Assessment of Performance Unit mathematics test (APU, 1981), and this was used as an external criterion for the school mathematics test.

The questionnaire was administered by the two authors in withdrawal groups of 4-5 outside their usual mathematics classes. Pupils were informed about the purpose of the questionnaire and were given instruction on the use of the scales in the questionnaire. They were also assured of the confidentiality of their responses. This form of assessment enabled any queries about the questionnaire, its rationale and wording to be dealt with. The mathematics tests were administered by the mathematics teachers during normal class time.

There were difficulties in following up the initial sample nine months later owing to staff turnover in the school. This reduced the final sample size for the full analysis to 70; most of the missing subject data were from the lower ability sets.

Measures

To design the questionnaire, following Ajzen and Fishbein's (1980) methodology, a sample of 12 pupils were interviewed about learning mathematics in school. The interview was designed to elicit the following information: (1) their salient beliefs about the outcomes of learning mathematics in school; (2) their significant others with respect to learning mathematics, and (3) the kinds of behaviours in class that in their opinion would enhance or impede their learning mathematics. Several mathematics teachers were also interviewed about the kinds of classroom behaviours they thought might enhance or impede mathematics learning.

Following the elicitation phase, a questionnaire was designed on the basis of pupils' and teachers' responses to the interview that included the following variables:

(1) A *direct* measure of *attitude to learning mathematics* in school. Pupils evaluated the statement "learning maths in school" on four 5-point bipolar evaluative scales (useful-useless, bad-good, silly-clever, pleasant-unpleasant), and the mean evaluative rating for each pupil was used as the direct measure of attitude.

(2) An *indirect* measure of *attitude to learning mathematics* in school based on beliefs about nine possible outcomes of learning mathematics identified in the interviews and their evaluations of these outcomes (e.g., would help me get a good job when I leave school, would help me work out money questions when buying things). Beliefs were measured on a 5-point agree-disagree scale, and evaluative ratings (how good or bad the outcomes were) were also provided on a 5-point

scale. To obtain the belief-based measure of attitude, the belief rating with respect to each outcome was multiplied by its corresponding evaluation, and these products were then summed.

(3) *Subjective norm.* Pupils first indicated the extent of their agreement or disagreement with a set of normative beliefs concerning the expectations of important others identified in the interviews, such as parents, brothers/sisters, friends, teacher (e.g., "My parents think I should learn maths in school"). They then indicated their motivation to comply with each of these significant others (e.g., "I usually do what my parents think I should do"), again on a 5-point agree-disagree scale. Each normative belief was multiplied by its corresponding motivation to comply, and the sum of these products served as the measure of subjective norm.

(4) *Behavioural intention.* From the interviews with pupils and teachers 12 learning behaviours that enhanced or impeded learning mathematics were identified (e.g., "listen to what the teacher is teaching", "ask the teacher sensible questions", "talk and distract others if I don't understand the maths"). Pupils indicated on a 5-point agree-disagree scale whether or not they intended to engage in each of these behaviours over the following two terms.

(5) *Learning behaviours.* Two measures of pupils' learning behaviours were obtained, the pupils' self-reported behaviour and the teachers' reports of their behaviour. Both pupils and teachers indicated on a 5-point scale how often the pupil had engaged in each of the 12 learning behaviours over the previous four weeks.

RESULTS

Reliability of measures

Cronbach alpha was calculated for the measures which were to be used in the multiple regression analysis. As Table 1 indicates these reliability indices were at a satisfactory level except for the composite subjective norm.

TABLE 1
CRONBACH ALPHA FOR MEASURES USED IN REGRESSION ANALYSIS

<i>Cronbach alpha</i>		
Composite attitude	0.76	(N = 130)
Composite subjective norm	0.32	(N = 142)
Learning behaviour intention	0.88	(N = 135)
Learning behaviour (self report)		
time 1	0.84	(N = 134)
time 2	0.81	(N = 106)
Learning behaviour (teacher report)		
time 1	0.93	(N = 141)
time 2	0.94	(N = 78)

It was found that one of the subjective norm items was unrelated to any of the other seven items ("Brother/sisters think I should learn maths"). When this item was removed from the set the alpha level increased to 0.67. The revised six-item measure of subjective norm was used in the subsequent analysis.

It was decided to use the composite attitude and subjective norm measures in the regression analysis instead of the direct measures. In the case of the direct measure of attitude the distribution was skewed, with low variability. It correlated at the 0.49 level ($N = 142$) with the composite measure. In the case of the direct measure of subjective norm the one item used presented some difficulties to the subjects. This was confirmed by the low correlation with the composite measure, $r = 0.14$ ($N = 142$).

The school-based mathematics tests were standardised to enable comparison of scores between the two year groups. The correlation of these initial standardised mathematics scores with scores on the independent APU test was 0.66 ($P < 0.01$, $N = 121$).

The correlation of teacher- and self-reported learning behaviours at both times was statistically significant at $P < 0.01$ level but indicated a low degree of common variance (see Table 2). This contrasts with the moderate retest correlations for the two modes of assessment.

TABLE 2

CORRELATIONS BETWEEN SELF- AND TEACHER- REPORTED LEARNING BEHAVIOUR

	<i>r</i>	<i>N</i>
Self — teacher report (time 1)	0.32	140
Self — teacher report (time 2)	0.29	73
Self (time 1) — self (time 2)	0.65	115
Teacher (time 1) — teacher (time 2)	0.63	79

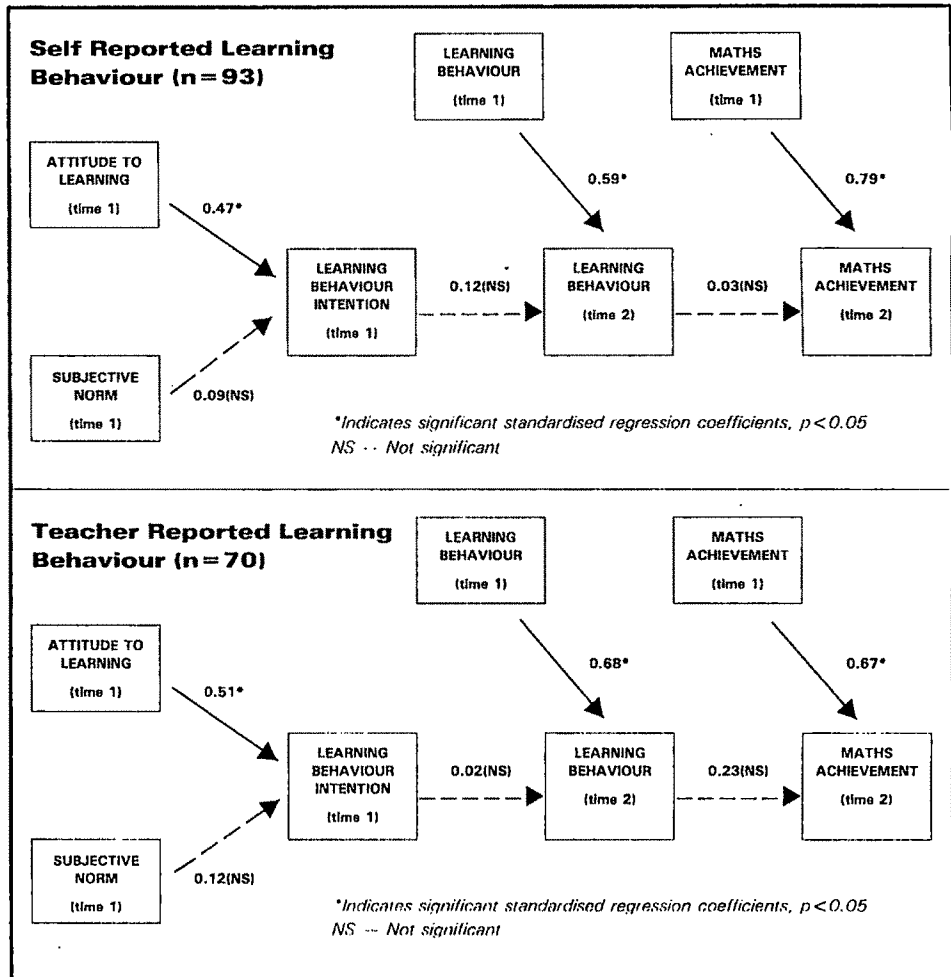
(All significant at $P < 0.05$)

Multiple regression analysis

Because of the low relationship between the self- and teacher-reported measures, two separate analyses were conducted using the SPSSx package for each mode of assessing learning behaviour. To test the theoretical relationships in Figure 1 a path analysis was carried out in which each variable in the path to the mathematics achievement variable (time 2) was regressed on prior variables. As Figure 2 shows, mathematics achievement at time 1 was the best predictor of subsequent mathematics achievement for both analyses (multiple R was 0.81 for both analyses). For the teacher index of learning behaviour only, learning behaviour was an additional independent predictor of subsequent achievement. The only significant predictors of both measures of learning behaviour were the corresponding prior measures of behaviour and not the intention, attitude or subjective norm variables. Attitude to learning was the only significant predictor of learning behaviour intention.

FIGURE 2

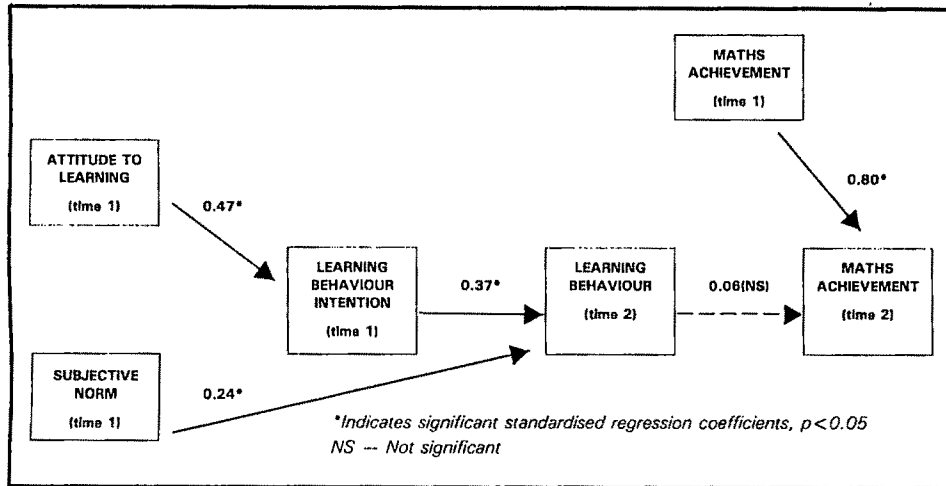
PATH ANALYSIS OF RELATIONSHIPS BETWEEN AFFECTIVE, BEHAVIOUR AND ACHIEVEMENT FACTORS



Behaviour intention was not a significant predictor of subsequent learning behaviour in either analyses. It was found, however, that if prior learning behaviour was not included in the analysis, then the pattern of relationships changed in the case of the self-report measure of learning behaviour (see Figure 3). Behaviour intention and subjective norm were both independent predictors of self-reported learning behaviour. That these relationships were not found when prior learning behaviour was included in the analysis is associated with the moderate correlations between the behaviour and affective variables at time 1. There was no corresponding predictive relationship between intention and behaviour as reported by teachers when prior behaviour was excluded from the analysis.

FIGURE 3

PATH ANALYSIS OF RELATIONSHIPS BETWEEN AFFECTIVE, BEHAVIOUR AND ACHIEVEMENT FACTORS
FOR SELF REPORTED LEARNING BEHAVIOUR
(initial learning behaviour not included)



In view of the low correlation between the two measures of learning behaviour and the different patterns of relationship between intention and behaviour when prior learning behaviour was excluded from the analysis it was decided to investigate the discrepancy between the two measures. It was decided to compare pupils whose self- and teacher-reported behaviours were similar with those whose self- and teacher-reported behaviours were different. The comparison was done in terms of their level of mathematics achievement, other affective variables, sex, age and mathematics ability set.

Pupils were sorted into three groups based on the relationship between their self- and teacher-reported learning behaviour:

Matchers — those whose self report was within one standard deviation above or below the teacher report;

Over-reporters — those whose self report was more than one standard deviation above the teacher report;

Under-reporters — those whose self report was more than one standard deviation below the teacher report.

For both assessment times, there was a significant linear relation between mathematics achievement level and whether pupils self-reported behaviour was under, matching or above that reported by teachers (see Table 3). Pupils who rated their behaviour well below their teachers' rating scored above average on the mathematics test, those who rated their behaviour above their teachers scored below average on both occasions.

TABLE 3

MEANS, STANDARD DEVIATIONS AND F VALUES FOR ACHIEVEMENT LEVELS OF PUPILS GROUPED BY MATCH OF TWO ESTIMATES OF LEARNING BEHAVIOUR

		Under- Reporters	Matchers	Over- Reporters	F- Values
Initial maths achievement level (standardised scores)	mean	0.54	0.18	-0.64	11.3*
	SD	0.9	0.9	1.0	df=2,129
	N	16	87	29	
Second maths achievement level (standardised scores)	mean	0.93	0.46	-0.15	5.8*
	SD	0.7	0.8	0.9	df=2,65
	N	8	42	18	

(* $P < 0.5$)

Cross-tabulating pupil rating group with ability set membership at both assessment times showed a significant relationship (see Table 4). For pupils whose self report matched their teacher's report most were in the top set, with decreasing proportions in the lower sets 3 and 5 — 42 per cent, 32 per cent and 26 per cent respectively. The corresponding proportions for the second assessment were 68 per cent, 21 per cent and 11 per cent. Finally, there was a tendency for girls at initial assessment only to have a higher frequency of matching self reports than boys (chi-squared = 13.6, $df = 2$, $P < 0.05$). No other significant associations were found.

TABLE 4

CROSS-TABULATIONS OF MATHEMATICS ABILITY SET WITH LEARNING BEHAVIOUR GROUPS

time 1	Learning behaviour groups			
	Under- Reporters	Matchers	Over-Reporters	
maths ability set				
1	6	39	5	50
3	7	30	10	47
5	4	24	16	44
	17	93	31	141
time 2	(Chi-squared = 10.1, $df = 4$, $P < 0.05$)			
1	7	32	6	45
3	1	10	3	14
5	0	5	9	14
	8	47	18	73

(Chi-squared = 16.1, $df = 4$, $P < 0.05$)

DISCUSSION

A Fishbein framework was adopted and modified to study and explain how affective factors might relate to school achievement, rather than merely establishing whether or not such relationships exist. The results have some bearing both on the theoretical framework and on some of the processes of learning mathematics in school. However, any conclusions have to take account of the methodological features of the study which took place in a busy comprehensive school in which large numbers of pupils and several teachers needed to collaborate. Withdrawal of pupils for administering the questionnaire posed no difficulties, though it was very time-consuming. Administering the mathematics tests and collecting the teacher reports on learning behaviour were less directly under the control of the authors. Despite efforts to follow up all of the initial sample and given the considerable number of variables in the study, there was a loss of data which was mainly from the lower mathematics ability sets. Both the sample size and the fact that the study was in one school reduces the generalisability of these findings. On the positive side, there were strong indications that pupils completed the questionnaires in a serious frame of mind and many appeared to find it an engaging experience.

Considering Fishbein's ideas about the determinants of behaviour intention, there was evidence that pupils with more positive attitudes, as measured by the composite expectancy-value procedure, reported more intention to engage in future mathematics learning behaviours. The evidence about subjective norm — perceptions that significant others think they should learn mathematics — did not fit the theory. Though correlated with behaviour intention, subjective norm was also correlated with the composite attitude measure and therefore had no independent association in the regression analysis.

One of the interesting results of this study is the lack of evidence for an independent predictive relationship between learning behaviour intentions and future learning behaviours for both measures of learning behaviour. The analysis indicates that this can be attributed to the association between the initial affective variables and the initial learning behaviour variable. It is the behaviour variable which is the only predictor of future learning behaviour. This evidence counts against the Fishbein theory and leads to a reconsideration of the notion of intention in the context of repeated multiple behaviours, such as those involved in school learning. There can be some doubt about the relevance of planning to do something which is already done repeatedly. Intentionality may be more relevant to the initiation of some new or different course of action. Attention might, perhaps, be better focused on plans to behave differently from current behaviour. It is also worth considering to what extent the order of asking about past behaviour and future plans can affect the assessment of behaviour.

Another notable finding was the independent relationship between teacher-reported learning behaviour and subsequent mathematics achievement. Though initial mathematics achievement was the best predictor of subsequent achievement in the case of teacher-reported learning behaviour, there was also an additional independent contribution from learning behaviour. That there was no such relationship with the self-report measure of learning behaviour can be linked to the low correlation between the two measures of learning behaviour. It is also associated with the predictive relationship between behaviour intention and future self-reported learning behaviour, when prior learning behaviour was excluded from the analysis. One side of the picture which emerged was that there is a Fishbein pattern of relationships between attitude, intention and self-reported behaviour. However, these self-report measures are not predictive of

learning behaviour if initial learning behaviour is taken into account. The other side of the picture is that teacher-reported learning behaviour by contrast with the self-report measure does predict a small proportion of the variance in mathematics achievement over and above that attributable to initial mathematics levels. This contrast between the two measures of learning behaviour in terms of their relationships to other variables and the low correlation between them suggests that the construct of learning behaviour could be the focus of any future study.

A third way of assessing learning behaviour, such as direct classroom behaviour, is needed to make sense of the different patterns of relationship and the weak link between the two measures of reported behaviour. The analysis of the mathematics achievement of those pupils who showed differences between self and teacher measures indicated that higher mathematics achievement was related to teacher reports being higher than self reports of learning behaviour. This could be interpreted as indicating that teacher reports were connected with, or perhaps, influenced by teachers' perceptions of mathematics achievement. Another analysis indicated that pupils who reported that their learning behaviour was well above that reported by teachers tended to be found in the lower ability sets. These findings point to the importance of situational and perceptual factors in the assessment of learning behaviour itself.

This study is part of an attempt to understand how affective variables might influence school learning and achievement. It has been based on an attempt to link the attitude and school learning literatures and has raised some further questions for investigation. Though the findings are not clear cut, they can be seen to illustrate the usefulness of considering more sophisticated affective constructs and processes in the school learning field. The study also reflects back on the Ajzen and Fishbein theory of reasoned action. It demonstrated the relationship between an expectancy-value interpretation of attitudes to particular behaviours and behaviour intentions in a field not previously studied. It also illustrated how prior behaviour can predict future behaviour without any significant additional contribution from prior attitudes, perceived prescriptions or intentions. This raises questions about the construct of behaviour intention in understanding behaviour of a repeated and multiple nature which takes place over an extended period of time, such as that found in a classroom. The realities of classroom life also lead to some questioning about whether there are other less formal behaviour intentions in secondary school classrooms which might be in competition with the formal school approved behaviour intentions of learning curriculum subjects. The presence of these informal "hidden" behaviour intentions could be considered from this framework but not without further adaptations.

Correspondence and requests for reprints should be addressed to Dr Brahm Norwich, Department of Educational Psychology and Special Needs, Institute of Education, University of London, 24-27 Woburn Square, London WC1H 0AA.

REFERENCES

- AJZEN, I., and FISHBEIN, M. (1980). *Understanding Attitudes and Predicting Social Attitudes*. Englewood Cliffs, N.J.: Prentice Hall.
- ASSESSMENT OF PERFORMANCE UNIT (1981). *Mathematical Development. Secondary School Survey Report No. 2*. London: HMSO.
- BENTLER, P. M., and SPECKART, G. (1979). Models of attitude and behaviour relations. *Psychol. Rev.*, 86, 452-464.
- BLOOM, B. (1976). *Human Characteristics and School Learning*. New York: McGraw Hill.
- BUTLER, R. (1986). The role of generalised expectancies in determining causal attributions for success and failure in two social classes. *Br. J. educ. Psychol.*, 56, 51-64.

- CLIFFORD, M. M. (1986). The effects of ability, strategy and effort attributions for educational, business and athletic failure. *Br. J. educ. Psychol.*, 56, 169-179.
- FREDRICKS, A. J., and DOSSETT, D. L. (1983). Attitude-behaviour relations: a comparison of the Fishbein-Ajzen and the Bentler-Speckart models. *J. Pers. soc. Psychol.*, 45, 501-512.
- HAERTEL, G. P., WALBERG, H. J., and WEINSTEIN, T. (1983). Psychological models of educational performance: a theoretical synthesis of construct. *Rev. educ. Res.*, 53, 75-91.
- MANSTEAD, A. S. R., PLEVIN, C. E., and SMART, J. L. (1984). Predicting mothers' choice of infant feeding method. *Br. J. soc. Psychol.*, 23, 223-231.
- SALTZER, E. B. (1982). The relationship of personal efficacy beliefs to behaviour. *Br. J. soc. Psychol.*, 21, 213-221.
- SCHIBECI, R. A. (1984). Attitudes in science: an update. *Stud. Sci. Educ.*, 11, 26-59.
- VOLLMER, F. (1986). The relationship between expectancy and academic achievement — how can it be explained? *Br. J. educ. Psychol.*, 56, 64-75.
- WEINER, B. (1986). *An Attributional Theory of Motivation and Emotion*. New York: Springer Verlag.

(Manuscript received 28th June, 1988)

PUPILS' PERCEPTIONS OF SCHOOL AND TEACHERS I - IDENTIFYING THE UNDERLYING DIMENSIONS.

By NOEL ENTWISTLE

(University of Edinburgh)

BELA KOZEKI

(Hungarian Academy of Sciences, Budapest)

AND HILARY TAIT

(University of Edinburgh)

SUMMARY. Previous comparative studies have shown interesting differences in motivation and approaches to learning between Britain and Hungary, which were considered to reflect different methods of teaching. The present study reports the development of scales designed to measure pupils' perceptions of school and teachers with the intention of relating those perceptions to school motivation and approaches to learning. The scales covered a wide range of aspects relating both to pupils' perceptions of school ethos and aspects of the learning environment. An inventory made up of 18 five-item Likert scales describing perceptions of school and teacher was given to samples of 516 12-15 year-old pupils in five British schools and a comparable sample of 602 pupils in Hungary. The factor structure of the school and teacher perceptions scales was almost identical in the two countries, suggesting that pupils perceive their schools in very similar ways in spite of the contrasting educational and social systems. There were differences between schools which, although small at the scale level, were sometimes large at the level of individual items. It is suggested that a revised set of scales which have been derived from these analyses might be used by schools to judge the way they are perceived by their pupils, and so also by the parents who will be having more influence in future on school policy and management.

INTRODUCTION

In a series of previous studies, comparisons have been made between pupils' school motivation and approaches to learning in Britain and Hungary (Kozeki and Entwistle, 1984; Entwistle and Kozeki, 1985). The current study extends this work, first by modifying the scales previously used, and then by developing a new series of scales designed to measure pupils' perceptions of school and of their teachers in secondary schools.

The most interesting findings from the previous comparative studies were, first of all, that the main dimensions of school motivation and approaches to learning could be identified equally clearly in both Britain and Hungary. Then, apart from differences in motivational patterns, there were also very significant differences in approaches to learning — the British pupils adopted stronger surface approaches and serialist styles, while the Hungarian pupils were more likely to endorse items relating to deep approaches and holist styles. In these studies pupils had not been asked about their learning environments, and thus the next step in the research involved extending our measurements in this direction. While perceptions of teaching were clearly of importance, it was recognised that a broader definition of learning environment would be necessary to include the school as a whole.

The main intention in developing the new scales was thus to investigate aspects of school ethos and school climate, together with perceptions of teaching,

to discover possible influences of these on pupil motivation and approach to learning. Such influences have already been demonstrated in higher education (Entwistle and Ramsden, 1983), and similar patterns were anticipated, at least in secondary school. It would have been possible to draw on literature on school effects and school climate (Cuttance, 1986; Finlayson, 1987), but the variables measured in such studies describe the "macro-level" of school organisation and the social composition of pupil intakes. Our concern was at the level of interaction between pupils and teachers and of the *qualitative* effects that learning environments are expected to have on pupils' learning, which involved a closer reliance on the parallel literature on student learning. This focus led to a rather different set of variables which could, ultimately, be used to complement the macro-level analyses.

In the research in higher education, relationships have been discovered between student perceptions of their main department and their approaches to studying, but not their levels of motivation or their work habits (Entwistle and Ramsden, 1983). Those relationships were revealed mainly at the departmental level, by correlating mean scores of students in a substantial number of departments. Although this, again, is the ultimate aim of the current research, the first step has been to investigate whether secondary school pupils are able to make equivalent judgments. Thus this study has used a small sample of schools in each country, selected to ensure substantial differences between them. If it proved possible to demonstrate differences in pupils' perceptions of those schools, then it would be worth moving on to a larger-scale study of what might be called "school and teacher effects".

CONCEPTUALISATION AND SCALE DEVELOPMENT

Perceptions of school and teaching

Whereas most of the scales developed previously to measure motivation and approaches to learning had been used extensively in previous studies at school or university level, the scales used to describe perceptions of school and teachers were less securely rooted in the previous literature. It seemed essential to draw on what was already known about the effects of university departments on approaches to learning to inform the development of scales describing pupils' perceptions of school and teachers.

A heuristic model of the teaching-learning process in higher education had already been developed and extrapolated to teaching and learning in schools (Entwistle, 1987a, 1987b) and this was used to guide the selection of scales. Also, at student level, Ramsden (1981) had developed a *Course Perceptions Questionnaire*. Only three of the subscales on that questionnaire had been found to relate to approaches to learning, irrespective of subject area. Deep approach was associated with perceived "freedom in learning" and "good teaching", while surface approach was linked with a lack of "freedom in learning" and a heavy "workload". In interviews, students repeatedly mentioned the influence of assessment procedures and also indicated what they perceived to be the components of "good teaching" as they related to lectures — level, pace, structure, explanation, enthusiasm and empathy (Entwistle, 1987b).

The next step in this study involved finding aspects of the learning environments in secondary schools which would parallel those found at university in influencing motivation and approaches to learning. It was also necessary to explore pupils' perceptions of the school which, presumably, would involve those

all too elusive concepts of "school ethos" and "school climate". Some attempts have been made to define these concepts, but little in the literature was found specifically to guide the creation of scales of pupils' perceptions.

The perceived "aims of the school" were considered to be one area which might help to describe school ethos, combined with perceptions of "relevance", "friendliness" (or social climate) and "discipline". (Defining items of these and the other scales describing perceptions of school and teachers will be found in Table 1.)

TABLE 1
SCALES AND DEFINING ITEMS OF PERCEPTIONS OF SCHOOLS AND TEACHERS

Scale	(Cronbach alpha)	Defining Items
<i>School Ethos</i>		
Aims of School		(see Table 5)
Friendliness	(0·60)	Most of the pupils in the class get on well together
Relevance (neg.)	(0·72)	This school doesn't seem to provide much knowledge which will be useful in later life
Discipline	(0·70)	There is really rather little bad behaviour in this school
<i>Learning Environment</i>		
Formality	(0·53)	Our teachers seem to spend a lot of the lesson talking, without letting us join in
Workload	(0·73)	We are given far too much work to do in this school
Factual Assessment	(0·52)	Too many teachers ask us questions in class just about facts
Openness	(0·50)	Our teachers seem interested in what pupils have to say
Facilitating Learning	(0·77)	Our teachers explain to us how to go about studying
<i>Teaching Effectiveness</i>		
Explaining	(0·69)	Our teachers are generally good at explaining things to us
Simplifying	(0·71)	The notes that most of our teachers give us are clear and useful
Organising	(0·70)	Most of our teachers present their lessons in a well- organised way
Holist (neg.)	(0·69)	Too many teachers wander off the point so we can't follow them
Serialist (neg.)	(0·65)	Too many teachers give us endless facts and details
<i>Teacher-Pupil Relationships</i>		
Enthusiasm	(0·68)	A lot of our teachers really seem to enjoy what they are teaching
Support	(0·74)	Teachers here are always ready to listen to our problems
Control	(0·53)	Our teachers keep a close eye on what we have done for homework
Criticism	(0·66)	We need more praise and encouragement from most of our teachers

Scales indicating perceptions of "workload", "openness" (rather than freedom in learning which is a characteristic more of higher education), "formality" in teaching, and "factual assessment" were derived directly from the university-level research. The heuristic model suggested the inclusion of study skills support, phrased to indicate efforts at facilitating learning, and also several of the components relating to the perceptions of good teaching. Students had perceived the

importance of pitching lecture material at the right level, and this was translated into a scale of "simplifying". Similarly "structure" became "organising", but "explaining" was left with the same label. To parallel the pupils' own learning styles, scales of "holist" and "serialist teaching" were developed, intended to indicate when teachers were perceived to be using too extreme a teaching style.

Students had mentioned the value they placed on lecturers' enthusiasm and empathy in helping them to learn. These were expanded to include four aspects of teacher-pupil relationships — "enthusiasm", "support", "control" and "criticism".

METHOD

Instruments

The scales described in the previous section were used as the basis of the second part of an inventory called *Pupils' Feelings about School and Schoolwork*: the first part is described in the subsequent article relating perceptions of school and teachers to motivation and approaches to learning. The inventory was translated into Hungarian, in semantic rather than simply linguistic terms, and then tried out with a small sample of pupils in each country to check on its intelligibility. As the English items had been jointly discussed as they were developed, equivalence of meanings could be to a large extent guaranteed.

Each of the 18 scales contained five Likert-type items with five response categories — definitely agree, agree to some extent, cannot decide or does not apply, disagree to some extent, and definitely disagree. The large overall number of items in the whole inventory (180) meant that it had to be divided into two parts for the purposes of administration. Part A contained items relating to motivation and approaches to studying, which will be discussed in the following article, while Part B included the scales described above.

As some of the scales, especially those in Part B, were being tried out for the first time, it was anticipated that several scales would require subsequent amendment or might prove altogether inadequate. Indeed the extent to which pupils would be able to rate aspects of teaching in a consistent manner was accepted to be problematic, and the range and variety of items were designed to allow those scales and items which proved effective to be selected and adapted for use in subsequent work. The Cronbach alpha values obtained with the current scales are shown in Table 1 and can be considered adequate for scales of this length and type. It is difficult to judge what level of consistency such scales "should" have, but in Part A of the inventory two five-point Likert scales, each consisting of five well-defined personality items, had Cronbach alpha coefficients of 0.62 and 0.66 respectively. It might be considered, therefore, that coefficients above, say, 0.5 are "acceptable", while those above 0.6 indicate strong homogeneity in this type of scale.

Sample

Five secondary schools in Edinburgh and its locality were selected to provide a wide divergence of types. Schools 1 and 2 were independent schools, the first of which was single sex (girls). School 3 was a comprehensive school with an academic tradition in a country town, while Schools 4 and 5 were city centre comprehensives, the first having a mixed intake and the other drawing from an area of poor housing, but with a reputation for having a particularly caring staff.

The comparable sample in Hungary was drawn from Budapest. Schools 1 and 2 had pupils in the 13 and 14-year age groups in the same building, while the remaining three schools were actually pairs of schools drawing from the same

catchment area. School 1 has a strong academic tradition with accompanying formality. School 4 is a new school in a workers' area of Pest which is considered to have a particularly good school "ethos".

In both countries, two age groups were selected, 12-13 years and 14-15 years, with two classes being chosen to represent each age group in each school. Clearly this method of sampling would not provide an adequate estimate of the population in each school, but it was appropriate for the purposes of this particular study.

The total sample contained 516 British pupils (322 girls and 194 boys; 253 aged 12-13 and 263 aged 14-15) and 602 Hungarian pupils (287 girls and 315 boys; 277 aged 12-13 and 325 aged 14-15).

Administration

Schools were asked to make available two teaching periods to allow each part of the inventory to be administered separately. Pupils were asked to write their names on each part of the inventory, but it was explained that the teacher would put the responses into an envelope and seal it. They were also told that no one in the school would see their responses. Identification of pupils was necessary in order to relate responses to teachers' ratings, but it was recognised that this procedure might encourage some pupils to make more favourable judgments than on an anonymous questionnaire.

RESULTS

Initially, item response distributions and item-scale-total correlations were calculated, as part of the scale development work for future studies and also to ensure that no items had been misallocated. No item was rejected by this procedure, although several weak items were identified. Scale totals for each scale were computed and used to explore the dimensionality of the inventory.

Mean scores were examined by school, and subsequently by gender and by age. As there was some doubt about the consistency of pupils' responses, item frequency distributions by school were also considered for a selection of items.

Dimensionality of the Inventory

Factor analysis by the maximum likelihood method were carried out using the SCSS program (Nie *et al.*, 1975) with varimax rotation to simple structure creating orthogonal factors. The decision to report the orthogonal structure was taken only after inspecting oblique solutions. As the structures were basically very similar, the greater interpretability of the orthogonal solution was preferred. The number of factors to be extracted was determined by the criterion of eigen values greater than one.

The dimensionality of the perceptions of school and teachers is shown in Table 2. ("Aims" was left out of this analysis, as it could not be considered, conceptually, to be a coherent scale.) In both countries two factors were produced with a very similar pattern of loadings. It appears that there is a clear-cut division between scales showing positive attitudes to school and to teachers and those which express negative feelings. To confirm this somewhat surprising two-factor solution, further analyses were carried out at item level. The vast majority of items, in both countries and for several different forms of factoring, fell repeatedly into the first two factors, with only groups of three or four items left in the subsequent factors. It seems, therefore, that only two factors can be established from the 90 items included.

Although this finding suggests a disappointing oversimplification of the dimensionality implied in devising the scales, it is closely similar to what was found with university students (Entwistle and Ramsden, 1983). In that study, one factor from the *Course Perceptions Questionnaire* described formal teaching methods linked to vocational relevance, while the second factor contained a cluster of positive ratings, related to good teaching, openness to students, freedom in learning, friendliness, and, to a less extent, clear goals and standards, and a lighter workload.

TABLE 2

FACTOR ANALYSIS OF PUPILS' PERCEPTIONS OF SCHOOL AND TEACHERS, BY COUNTRY

Scales	Factor Loadings	
	Britain	Hungary
<i>School Ethos</i>		
Friendliness	46	43
Irrelevance	-30	68
Discipline	52	65
<i>Learning Environment</i>		
Formality	64	72
Heavy Workload	72	64
Factual Assessment	78	72
Openness	72	64
Facilitating Learning	65	76
<i>Teaching Effectiveness</i>		
Explaining	75	83
Simplifying	76	82
Organising	79	83
Too Holist		77
Too Serialist	76	76
<i>Teacher-Pupil Relationships</i>		
Enthusiasm	73	74
Support	78	78
Control	55	52
Criticism	75	74

Two factors explained 53% of the variance in Britain and 54% in Hungary
 Decimal points and loadings below 0.30 have been omitted

In schools a very similar picture emerges but, with the increased number of scales phrased in a "negative" direction, a stronger "formal teaching" factor is found associated with heavy workload, factual assessment, criticism and both of the extreme teaching styles. Several aspects are of particular interest in Table 2. Above all, the similarity in the loadings in the two samples from Britain and Hungary, given the different educational and social systems, is remarkable. The first factor puts together the same string of "good teaching" variables, with control, discipline and friendliness forming a weaker triad in each country. But still, strong teacher control in the classroom is clearly seen by pupils in both countries to be part of effective teaching, while too heavy a workload, over-severe criticism and too factual an approach are all associated with formal teaching methods and are not appreciated by the pupils.

Differences between schools in Britain and Hungary

One of the main purposes of this study was to develop scales capable of discriminating between different schools. Thus one test of validity is to see whether such differences are found. It is also of interest to consider whether there are systematic differences between the two samples of pupils in Britain and Hungary in the ways they perceive their schools and teachers, although such cross-cultural comparisons with questionnaire instruments are fraught with difficulty.

TABLE 3
MEAN SCORES ON SELECTED SCALES BY SCHOOL AND COUNTRY

Scale	British Schools					Total Brit.	Hungarian Schools					Total Hung.
	1	2	3	4	5		1	2	3	4	5	
School Friendliness	20.5	21.1	19.4	19.4	19.6	20.2	18.9	19.9	20.3	19.6	19.0	19.5
School Irrelevance	11.3	10.4	13.9	15.1	16.1	12.8	13.2	11.9	11.8	11.9	12.1	12.3
School Discipline	18.3	19.2	15.9	15.8	16.9	17.6	14.8	15.6	16.7	17.7	16.3	16.1
Teacher Control	20.3	20.8	20.8	19.7	20.6	20.4	18.0	19.1	19.5	19.5	19.5	19.1
Teacher Support	19.2	19.6	18.5	18.6	19.5	19.2	16.9	18.5	19.2	19.4	17.7	18.3
Teacher Enthusiasm	18.5	18.5	17.9	18.1	19.3	18.5	17.4	19.0	19.5	19.7	19.0	18.9
Skill in Explaining	18.7	19.2	18.8	18.3	18.6	18.8	15.7	16.9	17.7	18.0	16.6	17.0
Facilitating Learning	18.3	16.1	17.9	16.9	19.0	17.7	15.6	16.4	18.1	18.4	17.6	17.2
Workload	13.5	12.7	15.5	15.8	16.3	14.4	16.2	14.7	15.7	16.9	15.7	15.8
Openness	18.7	19.7	18.3	18.2	19.2	18.9	17.3	18.3	18.6	18.7	18.3	18.2
Formality	17.6	16.5	18.6	17.3	18.1	17.6	15.5	16.0	15.4	16.1	16.2	15.8
Teacher Rating-Exam	3.5	4.6	2.9	3.4	3.2	3.6	3.7	3.6	3.4	3.3	3.5	3.5
(N)	(167)	(115)	(75)	(70)	(89)	(516)	(126)	(137)	(105)	(123)	(111)	(602)

*Indicates significantly higher mean for that country
(t-test, $P < 0.05$)

Table 3 represents the mean scores on a range of the scales, selected from the most effective ones, to cover all the factors identified in the previous section. The method of sampling makes detailed tests of statistical significance inappropriate, and differences between countries can be interpreted only in relation to subgroups which may themselves differ (see Table 4). To see whether the pupils' perceptions of school and teachers do seem to show interesting variations between schools, the patterns of variation across each group of five schools were examined.

The greatest number of differences between the schools in the sample was found in Hungary, although far and away the largest difference was found in Britain for school irrelevance. Large differences were found in both countries in school discipline, workload, and "facilitating learning". In these Hungarian schools, quite large differences emerged on three of the teacher scales, while the British schools in the sample differed somewhat more in perceived formality.

Gender and age differences in mean scores

Although this was not one of the main purposes of this study, Table 4 reports differences between boys and girls, and between the two age-groups used to make up the sample.

TABLE 4
MEAN SCORES IN SELECTED SCALES BY GENDER, AGE-GROUP AND COUNTRY

	Britain		Hungary		Britain		Hungary	
	Male	Female	Male	Female	12-13	14-15	12-13	14-15
School Friendliness	19.6	*20.5	19.3	19.8	20.2	20.1	19.6	19.4
School Irrelevance	13.7*	12.3	12.6	12.0	13.1	12.5	13.6*	11.2
School Discipline	16.9	*17.9	16.6*	15.7	18.1*	17.0	15.8	*16.5
Teacher Control	20.6	20.4	19.0	19.1	20.5	20.4	19.2	19.0
Teacher Support	18.9	19.3	18.3	18.3	19.5	18.9	18.6	18.1
Teacher Enthusiasm	18.2	18.7	18.5	*19.3	18.8	18.2	19.5*	18.4
Skill in Explaining	18.7	18.8	17.1	16.8	19.1*	18.4	17.5*	16.6
Facilitating Learning	17.3	17.9	17.6*	16.6	18.4*	17.0	18.2*	16.3
Workload	15.0*	14.0	16.2*	15.4	14.9*	13.9	15.6	16.0
Openness	18.7	18.9	18.0	18.5	19.1	18.6	18.1	18.3
Formality	17.7	17.5	16.1	15.6	17.3	17.8	16.3*	15.4
Teacher Rating-Exams	3.7	3.5	3.4	3.7	3.5	3.6	3.5	3.5
(N)	(194)	(322)	(315)	(287)	(253)	(263)	(277)	(325)

*Indicates mean of that sub-group significantly higher than its pair (t-test, $P < 0.05$)

The British girls in the sample perceive school as being more friendly, while British boys, and younger pupils in the Hungarian sample, see the school as being more irrelevant. There is a large difference between girls in British and Hungarian samples, with the British girls seeing it much more favourably. Younger pupils in Britain are more likely to report discipline as good than older pupils or than Hungarian pupils in the sample.

Hungarian girls and younger pupils are more likely to agree that their teachers are enthusiastic, while more younger than older pupils in both countries perceive their teachers as being good at explaining and facilitating learning. It is not at all clear how to interpret these lower ratings by older pupils, unless it represents a developing awareness of their need to understand more fully what they are learning. Certainly, the older pupils, at least in Britain, are beginning to organise their work habits more effectively by that time. Boys in both countries are more likely to perceive a heavier workload, while in this sample younger pupils in Hungary see their teachers as being more formal.

The differences in mean scores do not, however, form a sufficiently clear pattern to give confidence that the differences have any educational significance.

Analyses of individual items

As there has been a recent suggestion (Meyer, 1988) that perceptions of learning environment may be too specific to justify their inclusion in scales, it was decided to inspect a range of individual items as well. This possibility was explored first by examining the set of items described as "school aims", which were, in any case, not conceptually homogenous. Pupils' responses to these items are shown by school and country in Table 5, while Table 6 explores similar differences for a selection of other items. As it is impossible to present the whole distribution of responses to each item in a journal article the percentages of definite agreement and definite disagreement are reported here, together with the difference between the two, which can be used as an indication of overall differences between schools.

Table 5 records what seem to be exceptionally large differences, both between schools and between countries. Two state schools in Britain and one in Hungary show "friendly climate" as a strongly recognised aim, while all but one of the Hungarian schools and the two independent schools in Britain are seen to be emphasising "dependability". The differences in creative activities are less marked, but Hungarian pupils, in one school in particular, seem not to believe that the school is really trying to get the best out of all the pupils, while pupils in all but one of the British schools believe this to be an aim. Involvement with the local community varies a good deal with pupils in the least favourable social environment in Britain, and in the equivalent Hungarian school, rating this aim particularly strongly.

Many of the items showed substantial differences between schools, much larger, it would seem, than the differences seen in the comparisons of mean scores. Table 6 represents a selection of items, chosen to show the range of differences between schools and countries which were present at item level. In three cases, groups of three rather similar items have been chosen to indicate the extent to which the same, or different, patterns of response are shown across the ten schools.

Some of the large differences in the British schools can be attributed to social class and intellectual differences in the intakes of the schools. Thus the item "Most of us are here only because we have to be" follows exactly what would be a rank-ordering of the schools in terms of intake. It is, however, perhaps worth noting that some 42 per cent of the British pupils in the state schools definitely agreed with this item, compared with only 14 per cent in Hungary (although some of the British comprehensive schools will have been "creamed" of the most able pupils).

The two items describing teacher formality both show strong and consistent differences between countries. Hungarian teachers are seen by pupils in this sample as less likely to use dictation and encouraging copying from the board (31 per cent compared with 51 per cent) or to dominate the lessons through "teacher talk" (8/18 per cent). Taking these findings together with the perception that Hungarian teachers are more ready to make links with real life (31/19 per cent), the higher levels of deep approach and holist style found in previous studies, and in this one, become more understandable in terms of what may be real differences in teaching methods or style.

Differences in individual schools across both countries are also marked. Generally pupils in the British schools in this sample believed teachers were shown respect to a much greater extent than in Hungary, yet School 4 in each country strongly contradicts that general pattern.

TABLE 5
PERCENTAGE DEFINITE AGREEMENT AND DISAGREEMENT WITH PERCEIVED AIMS OF SCHOOL AND COUNTRY

Aim	British Schools						Hungarian Schools					
	1	2	3	4	5	Total	1	2	3	4	5	Total
Our school tries to promote a friendly climate among pupils and teachers	38	32	48	29	45	38	18	23	21	47	32	28
	2	2	1	4	5	3	7	2	2	2	3	3
	36	30	47	25	40	35	11	21	19	45	29	25
Our school puts a lot of emphasis on encouraging us to be reliable and dependable	36	42	31	14	29	32	23	55	49	61	47	47
	0	2	1	6	3	2	6	3	0	1	4	3
	36	40	30	8	26	30	17	52	49	60	43	44
In our school we are actively encouraged to become involved in creative activities	32	48	31	34	39	37	32	38	30	41	34	35
	0	4	4	3	5	3	4	4	4	3	6	4
	32	44	27	31	34	34	28	34	26	38	28	31
Our school really tries to get the best out of all its pupils	49	57	43	33	49	48	10	24	30	34	28	25
	1	0	1	6	6	2	10	4	8	2	7	6
	48	57	42	27	43	46	0	20	22	32	21	19
In our school we are actively encouraged to become involved with activities to support our local community	23	29	35	27	42	30	12	24	29	45	24	27
	7	11	8	0	7	7	9	4	4	1	5	5
	16	18	27	27	35	23	3	20	25	44	19	22
(N)	(167)	(115)	(75)	(70)	(89)	(516)	(126)	(137)	(105)	(123)	(111)	(602)

TABLE 6
PERCENTAGE DEFINITE AGREEMENT AND DISAGREEMENT WITH SELECTED PERCEPTIONS OF SCHOOL AND TEACHERS

Agree/Disagree		British Schools					Hungarian Schools						
Items (Scale)	(Difference)	1	2	3	4	5	Total	1	2	3	4	5	Total
A lot of my classmates are friends of mine (School Climate)		49/1 (48)	64/2 (62)	65/5 (60)	60/6 (54)	55/6 (49)	57/3 (54)	57/6 (51)	65/2 (63)	77/1 (76)	65/2 (63)	60/4 (56)	65/3 (62)
Most of us are here only because we have to be (School Irrelevance)		11/20 (-9)	10/38 (-28)	36/8 (28)	41/7 (34)	47/6 (41)	25/18 (7)	12/14 (-2)	9/24 (-15)	13/24 (-11)	19/29 (-10)	16/25 (-9)	14/23 (-9)
Generally there is a lot of respect shown to the teachers (School Discipline)		28/2 (26)	44/3 (41)	21/13 (8)	11/19 (-8)	28/9 (19)	28/7 (21)	17/10 (7)	19/6 (13)	19/3 (16)	34/3 (31)	21/6 (15)	22/6 (16)
When we are asked to read or do something out of school, most of our teachers check that it has been done (Teacher Control)		35/2 (33)	35/3 (32)	47/4 (43)	30/9 (21)	57/2 (55)	40/4 (36)	24/4 (20)	46/5 (41)	41/3 (38)	44/3 (41)	42/4 (38)	40/4 (36)
Most teachers here make a real effort to understand difficulties pupils may have with their work (Teacher Support)		26/2 (24)	37/6 (31)	41/3 (38)	49/6 (43)	36/5 (31)	36/4 (32)	16/7 (9)	34/4 (30)	31/3 (28)	35/6 (29)	32/5 (27)	30/5 (25)
A lot of our teachers seem to enjoy working with us (Teacher Enthusiasm)		26/2 (24)	35/6 (29)	31/7 (24)	24/9 (15)	51/0 (51)	33/4 (29)	10/15 (-5)	18/7 (11)	26/3 (23)	24/5 (19)	31/10 (21)	21/8 (13)
Many of our teachers seem really pleased when they see they've helped us to learn something (Teacher Enthusiasm)		19/4 (15)	21/9 (12)	24/1 (23)	30/3 (27)	43/2 (41)	26/4 (22)	15/6 (9)	37/4 (33)	30/4 (26)	31/3 (28)	36/4 (32)	30/4 (26)
Most of our teachers are ready to give us more of their time when we need it (Teacher Enthusiasm)		35/2 (33)	40/4 (36)	20/5 (15)	29/6 (23)	28/11 (17)	32/5 (27)	20/7 (13)	33/11 (22)	40/4 (36)	35/3 (32)	28/5 (23)	31/6 (25)
Our teachers generally help us to make links between different topics (Skill in Explaining)		15/5 (10)	21/8 (13)	19/7 (12)	19/6 (13)	27/11 (16)	19/7 (12)	14/7 (7)	36/4 (32)	41/2 (39)	40/4 (36)	23/8 (15)	31/5 (26)
Our teachers show us how to find what we need in books (Facilitating Learning)		17/5 (12)	13/11 (2)	28/4 (24)	17/4 (13)	43/7 (36)	22/7 (15)	6/14 (-8)	21/10 (11)	25/8 (17)	34/7 (27)	28/8 (20)	22/9 (13)
Our teachers seem interested in what we have to say (Openness)		22/2 (20)	41/3 (38)	25/5 (20)	26/9 (17)	36/5 (31)	30/4 (26)	20/7 (13)	34/4 (30)	22/3 (19)	36/4 (32)	28/4 (24)	28/4 (24)
We are given far too much work to do in this school (Workload)		4/14 (-10)	5/24 (-19)	16/25 (-9)	24/21 (3)	27/25 (2)	13/21 (-8)	10/20 (-10)	4/28 (-24)	7/12 (-5)	10/12 (-2)	8/19 (-11)	8/19 (-11)
Our teachers seem to spend a lot of the lesson talking about letting us join in (Teacher Formality)		12/10 (2)	7/15 (-8)	25/8 (17)	21/7 (14)	34/10 (24)	18/11 (7)	6/18 (-12)	7/29 (-22)	6/34 (-28)	8/33 (-25)	12/23 (-11)	8/27 (-19)
Many of our teachers dictate notes or make us copy from the board (Teacher Formality)		62/2 (60)	50/1 (49)	55/1 (54)	34/13 (21)	44/5 (39)	51/4 (47)	22/11 (11)	33/14 (19)	36/23 (13)	33/11 (22)	30/9 (21)	31/14 (17)
Too many teachers ask us questions in class just about facts (Factual Assessment)		4/8 (-4)	10/15 (-5)	24/8 (16)	21/1 (20)	32/12 (20)	15/10 (5)	6/14 (-8)	12/20 (-8)	17/13 (-7)	23/7 (4)	14/15 (-1)	14/15 (-1)
(N)		(167)	(115)	(75)	(70)	(89)	(516)	(126)	(137)	(105)	(123)	(111)	(602)

Looking at the patterns of response between schools at individual item level shows the value of this level of analysis. Although the scales seem quite homogeneous in terms of the Cronbach alpha coefficients, there are still interesting variations between items in some of the scales. For example, in "teacher enthusiasm" there may be real differences in how that enthusiasm is displayed in different schools. In the two British independent schools, for example, the teachers are perceived as putting somewhat less effort into their lesson preparation than teachers in the state schools, but to be more ready to give pupils additional individual attention. Of course, substantial consistency at individual item level cannot be expected with Likert scales, but the variations shown in Table 5 may point to a need to re-examine the roles of scales and individual items in describing school environments, as Meyer (1988) suggested.

DISCUSSION

Validity of school perception scales

What progress have we made towards being able to measure pupils' perceptions of schools and teachers? At one level, it seems reasonable to claim some success. There are substantial differences in sample means for several of the scales describing pupils' perceptions across schools and between countries, and those differences reflect at least those aspects of the schools which were known to the researchers. Thus, in the British sample, School 1 and School 4 both have a reputation for having a friendly relationship between staff and pupils, while School 2 has the strongest academic reputation of the five schools. School 5 is also known for its community involvement. Again, in the Hungarian sample, School 1, although successful in academic terms, does seem to be rather lacking in perceived friendliness, while the reverse is true of School 4.

There is a considerable difficulty in interpreting, at face value, pupils' perceptions of school and teachers, which stems from their differing social origins and attitudes. Thus, in the British sample in particular, some of the differences in perceptions of the school closely parallel school intakes. In particular, school irrelevance seems to be describing differences in the social origins and ability levels of pupils. Again the good teacher-pupil relationships in School 4 in Britain seem, perhaps, to have led those pupils, most of whom are by no means academically strong, to overestimate their own skills in learning quite substantially, and so such descriptions have to be interpreted with caution.

The differences between the samples drawn from Britain and Hungary also make sense in terms of what is already known about their educational practices in general. The degree of formality perceived by pupils in these British schools is clearly much higher than in Hungary, but so also is teacher respect. Worryingly, from a British point of view, school irrelevance is consistently higher in the British state schools in the sample. As mentioned before, perceptions of the teaching experienced in the two countries help to explain the differences in approaches to learning previously identified. At least among the individual items, the stress on relating academic knowledge to real world experiences was one of the most marked and consistent differences in favour of these particular Hungarian schools.

Refining the measurement of perceptions of school and teachers

Recent work by Ramsden and his colleagues (1989) has also sought to measure pupils' perceptions of their learning environments. Our approach is sufficiently different, with a much wider range of school perceptions being analysed, to suggest that both lines are worth continuing.

One of the intentions in using such a large number of items and scales in the inventory was to allow an improved, shorter version to be developed. This shorter version is now complete, and is described below to indicate the direction of future research, but it has yet to be used. Looking across the various analyses reported above, together with the total set of item-scale correlations, it was possible to identify certain additional aspects which should have been covered, while there were also some overlapping scales which could be merged to produce renamed composite scales.

The first step in improving the scales measuring perceptions of school and teachers was to remove the incoherent set of school aims. It was decided that the importance of aims in relation to school ethos would be better reflected by a separate, longer list which was not expected to produce any overall score. Such a list could also be given to teachers to allow comparisons to be made.

The next step was to remove ineffective scales. The teaching style scales had not worked in the way expected. Pupils responded to the different extreme styles in the same way — as general criticisms of the teaching rather than as contrasting perceptions of differing teaching styles. As there were a large number of items about teachers which merged together in the analyses, several other scales were dropped — skill in organising, skill in simplifying, factual assessment and teacher criticism. The remaining skill in teaching — explaining — was retained and strengthened as a major component of teaching effectiveness. This scale is now seen as complementary to “facilitating learning” which indicates another skill in helping pupils to develop more effective ways of learning. These two scales can be seen as defining two important aspects of teaching effectiveness, working in conjunction with the provision of an appropriate workload. It was decided to describe the “learning environment” of the school in terms of reworded versions of the scales of “formality” and “openness”, while an additional scale was created indicating how “well-organised” teachers are perceived to be.

Finally, scales of peer-group relationships have been added to complete the description of social relationships in the school. Scales of “companionship”, “collaboration”, and “rowdiness” have also been developed to parallel the three domains of motivation measured in the first half of the inventory — affective, cognitive and moral.

Conclusion

The set of scales in the revised form described above should prove valuable both for further research into the dimensionality of pupils' perceptions of school and the learning environment, and as a contribution to the debate about how to conceptualise school ethos. The sociological literature concentrates on defining school ethos in terms of the set of relationships observed in the school by visiting researchers. It seems that a full description of school ethos would have to reflect not only the observable facets of teacher and pupil behaviour, but also the complex of attitudes and perceptions which are an important part of any social organisation. The position of the headteacher and the management team may well be paramount, but the attitudes and perceptions of both teachers and pupils also need to be taken into account in conceptualising school ethos.

At a more practical level, it is believed that an inventory on pupils' perceptions could prove valuable for schools who are interested in their image in the community. Pupils' perceptions are the most direct influence on what parents come to

believe about the effectiveness of their local schools. As part of a scheme of self-appraisal at school level, the measurement of pupils' perceptions could provide additional valuable information for the consideration of the staff and governors.

From the point of view of the researcher, perhaps the most interesting question left unanswered by this study, so far, will be the extent to which pupils' perceptions of school and teachers are related to the pupils' own forms of motivation to learning. That question will be addressed in the next article.

Correspondence and requests for reprints should be addressed to Professor N. J. Entwistle, Department of Education, University of Edinburgh, 10 Buccleuch Place, Edinburgh EH8 9JT, Scotland.

REFERENCES

- CUTTANCE, P. F. (1986). *Effective Schooling*. A Report to the Scottish Education Department. Edinburgh: Centre for Educational Sociology.
- ENTWISTLE, N. J. (1987a). *Understanding Classroom Learning*. London: Hodder and Stoughton.
- ENTWISTLE, N. J. (1987b). A model of the teaching-learning process in higher education. In RICHARDSON, J. T. E., EYSENCK, M. W. and WARREN-PIPER, D. (Eds.), *Student Learning: Research in Education and Cognitive Psychology*. London: SRHE and Open University Press.
- ENTWISTLE, N. J., and KOZEKI, B. (1985). Relationships between school motivation, approaches to studying, and attainment among British and Hungarian adolescents. *Br. J. educ. Psychol.*, 55, 124-137.
- ENTWISTLE, N. J., and RAMSDEN, P. (1983). *Understanding Student Learning*. London: Croom Helm.
- FINLAYSON, D. S., (1987). School climate: an outmoded metaphor? *J. curric. Stud.*, 1, 163-173.
- KOZEKI, B., and ENTWISTLE, N. J. (1984). Identifying dimensions of school motivation in Britain and Hungary. *Br. J. educ. Psychol.*, 54, 306-319.
- MEYER, J. H. F. (1988). Student perceptions of learning context and approaches to studying. *S. Afric. J. higher Educ.*, 2, 73-82.
- NIE, N. J., HULL, C. H., JENKINS, J. G., STEINBRENNER, K., and BENT, D. H. (1975). *Interactive Version of the Statistical Package for the Social Sciences*. New York: McGraw Hill.
- RAMSDEN, P. (1981). A Study of the Relationship between Student Learning and its Academic Context. Unpublished Ph.D. thesis, University of Lancaster.
- RAMSDEN, P., MARTIN, E., and BOWDEN, J. (1989). School environment and sixth form pupils' approaches to learning. *Br. J. educ. Psychol.*, 59, 129-142.

(Manuscript received 10th October, 1988)

PUPILS' PERCEPTIONS OF SCHOOL AND TEACHERS II - RELATIONSHIPS WITH MOTIVATION AND APPROACHES TO LEARNING

BY NOEL ENTWISTLE

(University of Edinburgh)

BELA KOZEKI

(Hungarian Academy of Sciences, Budapest)

AND HILARY TAIT

(University of Edinburgh)

SUMMARY. The previous article has described the development of scales designed to measure pupils' perceptions of school and teachers. Previous comparative studies have shown interesting differences in motivation and approaches to learning between Britain and Hungary, which were considered to reflect different methods of teaching. This article explores the relationships between a set of inventory scores describing perceptions of school and teachers and another set indicating school motivation and approaches to learning. The complete inventory was given to samples of 516 12-15 year-old pupils in five British schools and a comparable sample of 602 pupils in Hungary. The factor structure of the combined inventory was investigated, together with correlational analyses at scale and item level which suggested that relationships did exist between perceptions of school and teachers, levels of school motivation, and approaches to learning.

INTRODUCTION

In the preceding article, the development of a series of scales was reported which represented facets of pupils' perceptions of school and teachers. The main purpose in developing these scales was to extend a series of previous studies in which comparisons had been made between pupils' school motivation and approaches to learning in Britain and Hungary (Kozeki and Entwistle, 1984; Entwistle and Kozeki, 1985). The idea, ultimately, will be to discover to what extent motivation and approaches to learning are affected by perceptions of school and teachers. Here the intention was more limited, namely to establish relationships between those two sets of variables.

The most interesting findings from the previous comparative studies were, first of all, that the main dimensions of school motivation and approaches to learning could be identified equally clearly in both Britain and Hungary. Then, apart from differences in motivational patterns, there were also very significant differences in approaches to learning — the British pupils adopted stronger surface approaches and serialist styles, while the Hungarian pupils were more likely to endorse items relating to deep approaches and holist styles. In these studies pupils had not been asked about their learning environments, and thus the next step in the research involved extending our measurements in this direction. While perceptions of teaching were clearly of importance, it was recognised that a broader definition of learning environment would be necessary to include the school as a whole. The scales reported in the preceding article were designed to cover that wider range of pupil perceptions.

In research in higher education, relationships have been discovered between student perceptions of their main department and their approaches to studying, but not their levels of motivation or their work habits (Entwistle and Ramsden, 1983). Those relationships were revealed mainly at the departmental level, by correlating mean scores of students in a substantial number of departments. Although this, again, is the ultimate aim of the current research, the first step has been to make initial analyses of relationships based on a relatively small number of schools.

CONCEPTUALISATION AND SCALE DEVELOPMENT

School motivation and approaches to learning

The earlier comparative studies had shown three main domains of school motivation — affective, cognitive and moral, and three main study orientations — meaning, reproducing and achieving. A weaker fourth “non-academic” orientation covered disorganised study methods and negative attitudes. A selection of the original subscales was made for the present investigation, but other scales were created to try to maximise connections between pupils’ perceptions of themselves and their school work, and the ways they saw their teachers and school.

Scales of affiliation, interest and responsibility were chosen to represent the motivational domains, while deep approach/holist style, surface approach/serialist style, and strategic approach were chosen from the study orientations. To supplement the measurement of pupils’ perceptions of their school work new scales of attitude to education, skill in learning, study skills and disorganised work habits were developed, drawing on equivalent scales used with university students (Entwistle and Wilson, 1977). (Defining items from each of these and the other scales will be found in Table 1.)

In the earlier inventory, parental involvement had been recognised only through the motivational dimension of “warmth”, but here this aspect was strengthened by including scales of parental support (similar to warmth) and parental control (a more positive version of what had previously been termed “adult pressure” — see Entwistle and Kozeki, 1985). To represent at least the negative aspect of peer-group influences, a scale of peer-group pressure was included.

Finally, to allow links to be made with earlier comparative studies (Eysenck *et al.*, 1980; Kozeki, 1988), short scales of extraversion and neuroticism, based on the Eysenckian concepts, were supplemented by an index of self-esteem, phrased negatively.

Form teachers were also asked to provide ratings on the pupils in their class on five-point scales covering achievement (related to likely level of performance in external examinations in Britain and the national rating system in Hungary which is already on a five-point scale), ability and effort. These ratings were included to allow an element of external validity to be introduced. Additional ratings of sociability and behaviour in class have not been used in the present analyses.

METHOD

Instruments

The 18 scales described in the section above were combined with 18 scales reported in the preceding article to form an inventory called *Pupils’ Feelings about School and Schoolwork*. The inventory was translated into Hungarian, in semantic rather than simply linguistic terms, and then tried out with a small sample of

TABLE 1
 SCALES AND DEFINING ITEMS OF PERCEPTIONS OF SELF AND SCHOOL WORK

Scale	(Cronbach alpha)	Defining Items
<i>Motivation</i>		
Affiliation	(0·55)	I really enjoy discussing with teachers ideas about life in general
Interest	(0·69)	I get very enthusiastic about some of my schoolwork
Responsibility	(0·57)	I always put a lot of effort into what we're asked to do at school
<i>Approaches and Styles</i>		
Deep	(0·63)	I often ask myself questions about the things I hear in lessons or read in books
Holist	(0·44)	I like to play around with ideas of my own, even if they don't get me very far
Strategic	(0·56)	I plan my working time carefully to make the best use of it
Surface	(0·44)	I find I have to rely on memorising a good deal of what we have to learn
Serialist	(0·37)	I'm very cautious about accepting what I read without having thought it through first
<i>Study Effectiveness</i>		
Attitude to Education	(0·50)	It seems to me that education should be mainly concerned with preparing us for adult life
Skill in Learning	(0·60)	I can usually pick out the important points in a lesson or in a book
Study Skills	(0·54)	I'm quite good at revising even a whole term's work
Disorganised Work Habits (negative)	(0·66)	I am easily distracted from my homework
<i>Home and Friends</i>		
Parental Support	(0·67)	My parents are always helpful and encouraging my schoolwork
Parental Control	(0·64)	My parents demand a lot of me and expect me to work hard
Peer-Group Pressure	(0·56)	It's important for me to keep in with my pals even if it means fooling around
<i>Personality</i>		
Extraversion	(0·62)	I find it easy to make friends
Neuroticism	(0·66)	I'm easily hurt if someone criticises me or my work
Low Self-Esteem (negative)	(0·56)	I often get discouraged at school

pupils in each country to check on its intelligibility. As the English items had been jointly discussed as they were developed, equivalence of meanings could be, to a large extent, guaranteed.

Each of the 36 scales contained five Likert-type items with five response categories — definitely agree, agree to some extent, cannot decide or does not apply, disagree to some extent, and definitely disagree. The large overall number of items in the whole inventory (180) meant that it had to be divided into two parts for the purposes of administration. Part A contained items relating to motivation and approaches to studying, while Part B included the scales described above.

The Cronbach alpha coefficients obtained with the Part A scales are shown in Table 1. All but three of the values are above 0·5 and these can be considered adequate for scales of this length and type.

Sample

Five contrasting secondary schools in Edinburgh and its locality were selected to provide a sample of 516 pupils. A comparable sample of 602 pupils in Hungary was drawn from five Budapest schools. In both countries, two age groups were selected, 12-13 years and 14-15 years with two classes being chosen to represent each age group in each school. (For further particulars of the samples refer to the preceding article.)

Administration

Schools were asked to make available two teaching periods to allow each part of the inventory to be administered separately. Pupils were asked to write their names on each part of the inventory, but it was explained that the teacher would put the responses into an envelope and seal it. They were also told that no one in the school would see their responses. Identification of pupils was necessary in order to relate responses to teachers' ratings, but it was recognised that this procedure might encourage some pupils to make more favourable judgments than on an anonymous questionnaire.

RESULTS

Dimensionality of the inventory

As the dimensionality of the school motivation and approaches to learning scales have been reported previously (Entwistle and Kozeki, 1985; Entwistle, 1988) and was similar in the current inventory, the results for Part A of the inventory will not be presented in full. Factor analyses were carried out using the maximum likelihood method on the SCSS program (Nie *et al.*, 1975) with varimax rotation to simple structure creating orthogonal factors. (See preceding article for a justification of this method.) In Britain, although five factors were produced, only two were substantial. The first combined the scales covering meaning and achieving orientations; they also included both parental scales and all three positive study method scales. The second factor covered reproducing and non-academic orientations with loadings on surface approach, disorganised work habits, peer group pressure, low self-esteem, and neuroticism. In Hungary, a similar structure was found but with separate factors for meaning and achieving orientations. Achieving orientation contained both parental scales together with responsibility and affiliation, while meaning orientation contained skill in learning, study skills and interest.

The dimensionality of Part B of the inventory was presented in the preceding article, and indicated that there were just two factors. The first of these represented positive evaluations of school and teachers, while the second consisted of items indicating negative feelings.

Factor analysis of the combined inventory was carried out with several scales removed. These were "aims" and its equivalent in Part A, "attitude to education", along with "skill in learning" (which can be considered as a self-rating of a criterion variable — school achievement), and the two Eysenckian personality dimensions (which tended to create separate small additional factors). The resulting factor analyses are shown in Table 2.

To a large extent, the factor analyses of the whole set of scales seem to do little more than repeat the two separate factor analyses carried out previously. The first two factors in each country are closely similar and represent the positive and negative evaluations of school and teachers. Factor III in Britain combines school motivation with meaning and achieving orientations, which are separated into

TABLE 2
FACTOR ANALYSIS OF THE COMPLETE INVENTORY

	I	II	Britain				I	II	Hungary			
			III	IV	V	VI			III	IV	V	VI
Affiliation			53		26				46	43		
Interest			67		30		29		38	47		
Responsibility	27		59				31		56	27		
Deep Approach			66						37	54		
Holist Style			54							58		
Strategic Approach			60	28			27		70			
Surface Approach		33		30				31			56	
Serialist Style			56				31		29		29	
Study Skills			62	-26						53		
Disorg. Work Habits		35		67				27	-34		69	
Parental Support	30		30						48			
Parental Control			46						42			
Peer-Group Pressure		44		40				36			37	
Low Self-Esteem		30		55				28			56	
Friendliness	42					26	36			25		
Irrelevance	-26	65				-42		71				
Discipline	44	-28	32				62					
Formality		64						71				
Heavy Workload		67		29				62				
Factual Assessment		74						72				
Openness	67						58			26		
Facilitating Learnng	64						73					
Explaining	68		33				80					35
Simplifying	72	-28					82					28
Organising	75						81					
Too Holist		76						77				
Too Serialist		81						73				
T. Enthusiasm	69		26				69					
T. Support	79						76	-26				
T. Control	46		34				49					
T. Criticism		76						72				

Six factors explained 51% of the variance in both Britain and Hungary.
Decimal points and loadings below 0.25 have been omitted.

Factors III and IV in Hungary. Factor IV in Britain, and Factor V in Hungary, represent the combined non-academic and reproducing orientations. The remaining factors are small and probably chance combinations, unless the two Factor VIs can be seen as a first indication of a separation between perceptions of school (friendliness and relevance) and teachers (enthusiasm and support).

Although the factors have mainly separated the two parts of the inventory, the rather slight remaining overlaps are still of interest. Taking both countries together, there is some indication of an association between school motivation and approaches to learning on the one hand, and pupils' perceptions of some aspects of their learning environment, on the other. In Britain, school motivation and the combined meaning and achieving orientation are linked with perceptions of good explanations given in an enthusiastic way, together with teacher

control and good discipline in the school. In Hungary, school motivation and meaning orientation are associated, albeit weakly, with perceived openness and friendliness. However, it must be stressed immediately that no causality, in either direction, can be inferred from such associations between the two parts of the inventory. Problems of interpretation will be discussed later.

Relationships between perceptions of school and pupil characteristics

One of the main purposes in planning this new direction in collaborative research was to explore possible interactions between school and teacher effects, as perceived by the pupils, and a range of individual difference measures. The current study can only begin this process in a tentative way, as there are too few schools to analyse the data at school level. Instead, relationships between the two halves of the inventory can be explored, both at scale and at individual item level. Which variables from the pupils' perceptions of their learning environments were related to their levels of motivation and approaches to learning? To what extent are the patterns of relationships in accord with what has been found in work with students?

TABLE 3
CORRELATION BETWEEN MOTIVATION/APPROACHES AND PERCEPTIONS OF SCHOOL/TEACHERS
BY COUNTRY

Scales Country	Deep		Strateg		Surface		Affil'n		Interest		Respons	
	Br	H	Br	H	Br	H	Br	H	Br	H	Br	H
School Friendliness	21	23	20	26	-01	06	27	34	20	23	21	27
School Irrelevance	-21	-17	-18	-12	23	26	-03	-14	-27	-27	-22	-22
School Discipline	29	28	31	33	-12	07	22	27	38	32	35	34
Teacher Control	29	29	31	19	07	13	20	29	28	20	39	26
Teacher Support	26	30	29	33	-12	10	20	33	35	37	34	42
Teacher Enthusiasm	32	29	30	33	-06	15	30	36	39	38	36	40
Skill in Explaining	41	38	38	36	-08	08	22	33	36	39	36	42
Facilitating Learning	18	28	31	33	04	15	18	29	23	32	20	34
Workload	-22	-07	-22	-16	25	24	04	-06	-18	-19	-17	-17
Formality	-01	-04	-07	-11	30	23	11	-04	-07	-16	02	-16
Teacher Rating- Exams	18	08	06	11	-25	-25	03	03	12	08	19	09
-Ability	16	05	-01	06	-25	-27	03	-02	10	08	12	05
-Effort												

Decimal points omitted

Correlations above 1.1 are significant in each sample

Note: for comparison purposes the highest correlations overall were 0.51 between deep approach and interest, and 0.50 between interest and responsibility.

Table 3 represents simple product-moment correlations between selections of the two sets of variables. Picking out the highest level of correlations in both columns and rows provides some indication of the patterns of relationships. Thus a deep approach is related most strongly, in both Britain and Hungary, to the teachers' perceived skill in explaining and, in Britain, it is also negatively related to workload at a level quite high for that particular variable. Strategic approach shows a substantial number of correlates. Besides skill in explaining, it is also related to "facilitating learning" and school discipline in the samples in both



countries, to teacher control and a lighter workload in Britain, and to teacher support and enthusiasm in Hungary. Surface approach is consistently related to school irrelevance, heavy workload and perceptions of formality in teaching, paralleling the negative evaluation factor described earlier.

Affiliation motivation is related, understandably, to friendliness and teacher enthusiasm, while interest is strongly related to positive evaluations of the teachers and a rejection of negative ones. Responsibility shows the largest number of substantial correlations, with the highest being on positive evaluations of teachers and on school discipline (with the addition of teacher control in Britain and facilitating learning in Hungary).

Regression analysis of this set of school and teacher perceptions against the six target scales from Part A produced multiple correlations varying from 0.40 to 0.54. Repeating this process at individual item level, together with an examination of item-scale correlations, indicated that the items shown in Table 4 were most closely, and consistently, related to the various scales of motivation and approaches to learning in the two countries.

TABLE 4

PERCEPTION ITEMS RELATED TO MOTIVATION AND APPROACHES

Deep Approach

Most of our teachers are good at showing how what we are learning helps us to understand the world outside. (Also associated with *Interest*)

Our teachers generally help us to make links between different topics and with real life.

Many of our teachers show us how to understand things better by asking the right questions. (Also with *Interest* and *Responsibility*)

Strategic Approach

We are generally given enough time to understand the things we have to learn.

A lot of our teachers are very good at showing us how to do the exercises or homework they set. (Also with *Interest* and *Responsibility*)

When new school rules are introduced they are usually followed. (Also with *Interest* and *Responsibility*)

Surface Approach

Teachers too frequently jump from one point to another, preventing us following what they're trying to say.

Our teachers seem more ready to see our mistakes than what we have done well.

Too few teachers test our understanding: they are more interested in what we have memorised.

Affiliation

A lot of teachers encourage us to make use of our own ideas. (Also with *Interest*)

Most of the pupils in this class are ready to help each other with their work.

Most of our teachers are good at encouraging even shy pupils to join in classroom discussions. (Also with *Interest* and *Responsibility*)

Interest

Most of our teachers show that they are interested in us as individuals.

Our teachers are generally good at explaining things to us. (Also with *Responsibility*)

Our teachers seem interested in what pupils have to say.

Responsibility

Our teachers set a high standard in what they expect of us.

Nearly all our teachers are ready to give us help and advice about our studying.

Our school really tries to get the best out of all its pupils.

Inevitably, items correlate with more than one scale and are included in several of the regression equations, so the three items used eventually to represent the relationships are indicative, rather than definitive. The criterion for choice was that the item appeared in both countries and showed a weaker relationship with other scales. This procedure inevitably left out items with fairly high correlations across several scales. The strongest of these remaining items were:

The rules in this school are generally sensible and fair
A lot of our teachers seem to enjoy working with us
Many of our teachers seem to put a lot of effort into preparing their teaching well
Most teachers here make a real effort to understand difficulties pupils have with their work

DISCUSSION

The ultimate goal of this research is to be able to identify the effects of schools on school motivation and approaches to learning. Differences in the perceptions of teaching in the samples in the two countries, reported in the previous article, suggested that British schools were perceived as being more formal with pupils being more respectful to teachers, while the Hungarian pupils found explanations related more to real-life experiences. Putting the relationships found in the analyses above into this context, there can be seen to be at least an indication that aspects of school ethos and learning environment, as perceived by the pupils, influence approaches to learning and school motivation.

The relationships which have been found are in line with previous work. Thus skill in explaining and a light work load are associated with higher scores on deep approach, while surface approaches are linked with heavy workload and formal teaching in both countries. Affiliation is related to friendliness and teacher enthusiasm and interest, predominantly to favourable perceptions of teacher-pupil relationships. Responsibility is again associated mainly with favourable evaluations of teaching. There is an interesting difference between the British and Hungarian samples in the relationships with teacher control. In Britain it is strongly related to responsibility and interest, while in Hungary it is more closely linked with affiliation. This finding almost certainly reflects a real difference in the ways control is predominantly exercised in the two countries — through formal rules Britain and through informal personal relationships in Hungary.

Of course, it can still be argued that all we have established is rather weak relationships between self-ratings of pupil characteristics and of their perceptions of school and teachers which could be inextricably interwoven in the pupils' responses. The reality of "school and teacher effects" can only be judged once a thorough analysis at school level becomes possible, but already in this study the evidence of differences in perceived teaching methods in Hungary linked to deeper approaches to learning represents important supportive evidence that we are dealing with more than just tautological statistical associations.

It is already possible to look to other emerging evidence of the reality of these relationships. In a study that has been carried out in parallel to the current study Ramsden *et al.* (1989) have been able to carry out in parallel to the current study pupil level and at school level with pupils in their final year of secondary education. The highest correlations they report at individual level are for deep approach with "independence in learning", achieving (or strategic) approach with "structure and cohesiveness", and surface approach with an "emphasis on formal academic achievement". Again, at individual level, it could be argued that

pupils hostile to the school ethos will rate school and their own study strategies in a less positive way without this implying any causal relationship, but Ramsden and his colleagues were able to show substantially higher correlations at school level across 50 schools.

Analyses of relationships between approaches and perceptions of the learning environment do, on occasions at least, produce insignificant results. Meyer and Parsons (1989) found no relationships in a study of students at the individual level. Of course, there are good reasons to expect much lower relationships at individual level. To the extent that students agree with a statement describing a particular aspect of a learning environment, the variance is substantially reduced and hence the correlation of that item with any other item will be low. Also it appears that aspects of the learning environment, such as assessment procedures, may affect all students uniformly, shifting the mean without changing the rank order. For example, introducing an essay-type examination increased the level of deep approach of all students on average, but left those originally adopting a deep approach still ahead of those who had previously adopted strong surface approaches (Thomas, 1986). Such influences could only be detected at the institutional level of analysis. Other aspects of the learning environment could, however, be expected to influence pupils in different ways, but still in ways which may be difficult to detect. For example, while a self-confident pupil might find almost any environment supportive, an anxious pupil might find the same environment threatening. This effect has already been demonstrated among students (Entwistle *et al.*, 1974). Thus at the very least, replicable effects may only be demonstrated if appropriately complex statistical analyses are adopted which allow variations and both pupil level and school level to be examined simultaneously.

Meyer and Parsons (1989) also challenge the existence of replicable factors in the items of the Course Perceptions Questionnaire and suggest the need to examine the effects of learning environments at the level of specific items, which can better identify qualitatively different aspects of the environments (Meyer, 1988). Our present study leaves this question open. The scales describing pupils' perceptions of school and teachers do not seem as strong as those describing school motivation and approaches to learning. It is not clear yet, however, whether this is a facet of real differences in the way learning environments should be described or simply a reflection of the more developed state of the conceptualisation of the scales of pupil differences.

A conceptual map

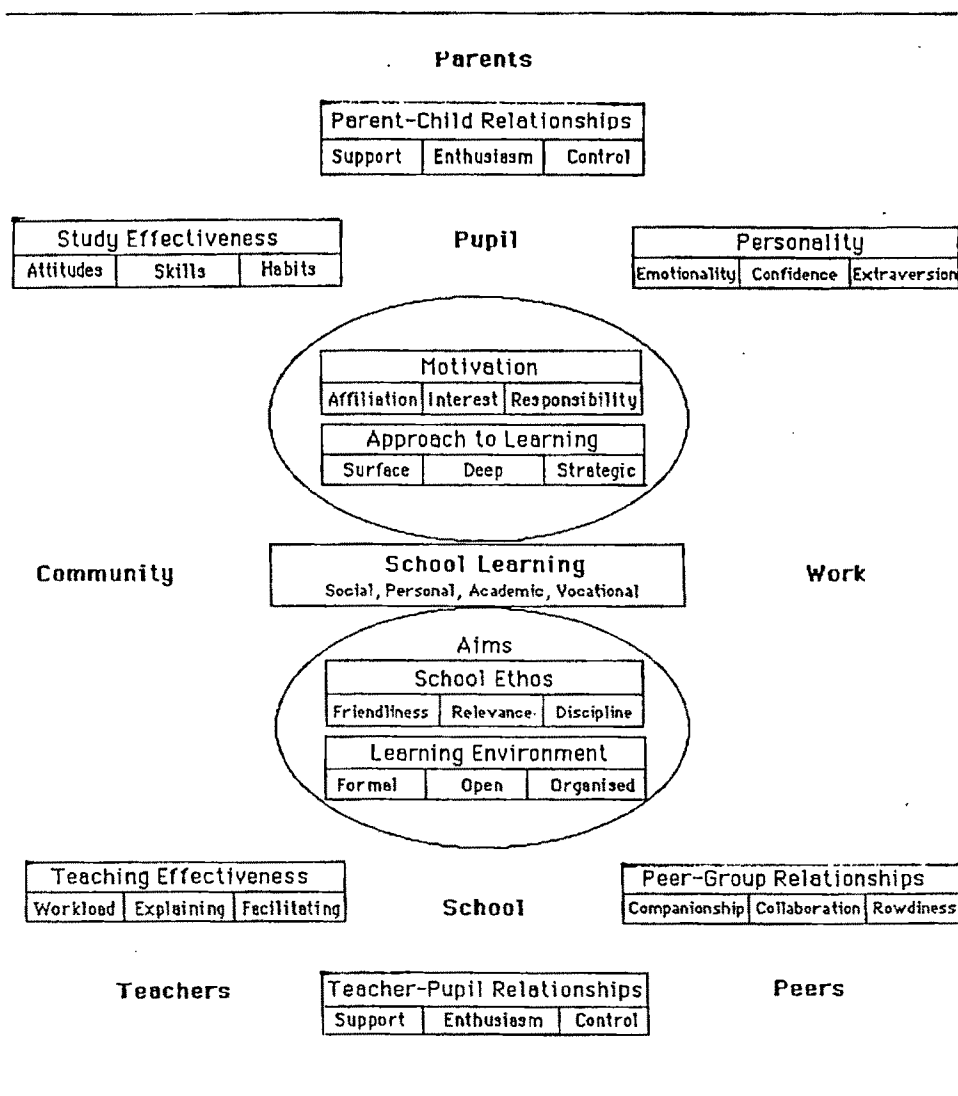
In the previous article, refinements of the scales covering pupils' perceptions of school and teachers were discussed. Changes in Part A of the inventory were also indicated by the analyses here.

The first step in improving this part of the inventory will be to remove the incoherent scale of "attitude to education" and to replace it with a scale measuring "instrumental attitudes" towards schoolwork. The factor analyses had indicated that the reproducing orientation was not being fully defined in the present instrument, hence the need to support it with a related scale.

Next it seemed better to reverse the direction of low self-esteem to produce "self-confidence", while a third, more cognitive aspect of parent-child relationships will be represented by a scale of "enthusiasm" for school-derived knowledge and skills. Finally, both styles of learning will be excluded, as both scales produced unsatisfactory Cronbach alpha values, perhaps suggesting that the items were difficult for secondary age pupils to answer.

FIGURE 1

A CONCEPT MAP INDICATING POSSIBLE INTERACTIONS BETWEEN PUPIL CHARACTERISTICS AND PERCEPTIONS OF SCHOOL



The whole set of scales can be seen in Figure 1 which maps the concepts seen, in the top half of the diagram, as the most important characteristics of the pupil and of parent-child relationships as they affect school learning. The bottom half of the figure represents pupils' perceptions of school and teachers, including peer-group relationships within the school. The concepts have been ordered within this concept map to reflect the underlying theories influencing the research strategy — the two triads of school motivations and approaches to learning and what

is expected to influence them. In each case the choice of scales and their position in the model has been influenced by one or other of these triads. Thus, for example, in the scales covering parent-child and teacher-pupil relationships, support is seen as affective, enthusiasm as cognitive, and control as moral, while in study effectiveness and teaching effectiveness instrumental attitudes and workload have been shown to induce surface approaches, study skills and good teaching to influence a deep approach, and work habits together with facilitating learning may be expected to be associated with a strategic approach.

Ultimately, complex statistical techniques, such as path analysis, will be required to tease out school and teacher effects on learning from those derived from child-rearing and individual differences between pupils. But such analyses, in themselves, are unlikely to be interpretable with any confidence due to the lack of a strong enough theoretical underpinning to make firm judgments about causality from the statistical models (Freedman, 1987). It will thus be necessary to supplement the multivariate analyses with further cross-cultural studies and also with detailed observations within individual schools. In this way a developing body of theory can be developed to provide firm help to teachers in conceptualising how the learning environment they provide will affect school learning of all forms — social, personal, academic and vocational. The model being developed is intended to be broad enough to encompass all the major educational aims and to encourage schools, and the educational system as a whole, towards a balanced emphasis on differing aims and a recognition that it is the quality of long-term learning, not simply the quantity of certificates collected, which is the ultimate criterion of the success of education.

Correspondence and requests for reprints should be addressed to Professor N. J. Entwistle, Department of Education, University of Edinburgh, 10 Buccleuch Place, Edinburgh EH8 9JT, Scotland.

REFERENCES

- ENTWISTLE, N. J. (1988). Motivational factors in students' approaches to learning. In SCHMECK, R. R. (Ed.), *Styles and Strategies of Learning*. New York: Plenum Press.
- ENTWISTLE, N. J., and KOZEKI, B. (1985). Relationships between school motivation, approaches to studying, and attainment among British and Hungarian adolescents. *Br. J. educ. Psychol.*, 55, 124-137.
- ENTWISTLE, N. J. and RAMSDEN, P. (1983). *Understanding Student Learning*. London: Croom Helm.
- ENTWISTLE, N. J., THOMPSON, J. B., and WILSON, J. D. (1974). Motivation and study habits. *Higher Educ.*, 3, 379-396.
- ENTWISTLE, N. J., and WILSON, J. D. (1977). *Degrees of Excellence*. London: Hodder and Stoughton.
- EYSENCK, S. B. G., KOZEKI, B. and KALMANCHEY, M. (1980). Cross-cultural comparison of personality: Hungarian children and English children. *Person. indiv. Diff.*, 1, 347-353.
- FREEDMAN, D. A. (1987). As others see us: a case study in path analysis. *J. educ. Stat.*, 12, 101-128.
- KOZEKI, B. (1988). Motivational and personality structures in education: individual differences in Hungarian school-aged children. In EYSENCK, S. B. G. and SAKLOFSKE, D. H. (Eds.), *Individual Differences in Children and Adolescents*. London: Hodder and Stoughton.
- KOZEKI, B. and ENTWISTLE, N. J. (1984). Identifying dimensions of school motivation in Britain and Hungary. *Br. J. educ. Psychol.*, 54, 306-319.
- MEYER, J. H. F. (1988). Student perceptions of learning context and approaches to studying. *S. Afric. J. higher Educ.*, 2, (in press).
- MEYER, J. H. F., and PARSONS, P. (1989). Approaches to studying and course perceptions using the Lancaster Inventory. *Studies in Higher Education*.
- NIE, N. J., HULL, C. H., JENKINS, J. G., STEINBRENNER, K. and BENT, D. H. (1975). *Interactive Version of the Statistical Package for the Social Sciences*. New York: McGraw Hill.
- RAMSDEN, P., MARTIN, E. and BOWDEN, J. (1989). School environment and sixth form pupils' approaches to learning. *Br. J. educ. Psychol.*, 59, 129-142.
- THOMAS, P. (1986). *The Structure and Stability of Learning Approaches*. Unpublished Ph.D. Thesis, University of Queensland, Australia.

(Manuscript received 10th October, 1988)

SEPARATE SKILLS OR GENERAL INTELLIGENCE: THE AUTONOMY OF HUMAN ABILITIES

By MICHAEL J. A. HOWE

(Department of Psychology, University of Exeter)

SUMMARY. That children's and adults' different intellectual abilities are to a considerable extent autonomous, and do not depend upon general ability or intelligence, is indicated by six kinds of evidence: (1) biographical and autobiographical reports, (2) patterns of ability in autistic and mentally handicapped individuals, including "idiots savants", (3) brain damage effects, (4) the absence of inter-task interference, (5) evidence of extraordinary abilities in people of average intelligence, and (6) findings suggesting that cognitive complexity may be unrelated to measured intelligence. It is argued that correlations between an individual's levels of performance at different abilities, which have often been interpreted as providing evidence of a general ability factor, can in fact be readily explained as being due either to elements that are shared by different tasks, or to any of a variety of a person's attributes that can influence performance at each of a number of tasks. Educational implications of acknowledging that separate intellectual skills are largely autonomous are discussed.

INTRODUCTION

TWENTY years ago, few would have disagreed with the view that human intellectual ability was to a considerable extent unitary and hierarchical, with some fundamental entity of general intelligence underlying (and constraining) individuals' differences in mental competence. It was assumed that knowing about a person's intelligence level served to help explain that individual's successes or failures at intellectual tasks. Debates about aspects of human intelligence had been in progress throughout much of the present century. They concerned, for instance, the contributions to a person's intelligence of general, group, and specific factors, the relative importance of hereditary and environmental influences, the uses and misuses of psychological testing, and the "fairness" of various tests. Nevertheless, there was widespread acceptance of the idea that people's intellectual achievements were constrained by a process of general ability. For some authors, this was synonymous with the factor "g", or the algebraic "positive manifold". The concept of intelligence was seen as having a necessary role in attempts to explain differences between individuals in their ability.

Today, such a view seems increasingly hard to defend. No one denies that the word "intelligence" is a useful one in everyday life, or that measures of intelligence have practical value, particularly in education. But the idea that information gained from administering intelligence tests contributes to an understanding of people's abilities has been questioned. Intelligence appears to be a purely descriptive concept, and not an explanatory one (Howe, 1988a, 1988b). Like other descriptive concepts, such as "productivity", it is useful for labelling and predictive purposes, but it cannot explain what it describes. Stating that someone performs well in certain situations because she is intelligent is no more meaningful than saying that a factory produces large amounts of goods "because it is productive". Also, some researchers have questioned the widely accepted viewpoint that intellectual skills are hierarchically controlled by some central process which affects all mental tasks in common. That view does not appear to

be at all firmly supported by empirical evidence. Contradicting the notion of central control, the different abilities of a child or an adult appear to be interconnected only to a limited extent (Rozin, 1976). The mental abilities of a single individual may form a rather loose, imperfectly connected, confederation of processing systems (Geschwind, 1983).

The present article demonstrates that there are strong empirical reasons for challenging the assumption that intelligence exists as a unitary entity. It argues that, contrary to that view, a person's different intellectual skills may be largely autonomous. They can be entirely independent of other abilities, and are unconstrained by any entity or quality of general intelligence.

EVIDENCE THAT ABILITIES ARE SEPARATE

Evidence from a number of sources supports the assertion that intellectual abilities are relatively specific, independent, and autonomous. First, there is a substantial amount of biographical data, much of which is frankly anecdotal. For example, there are numerous accounts of people of genius, such as Darwin, Freud, and Einstein, experiencing difficulty at tasks that many people find easy, and there are stories of brilliant chess players, musicians, and even mathematicians doing badly at skills outside their special expertise (Howe, 1982, 1989a; Gardner, 1988).

Secondly, there are numerous reports of mentally handicapped individuals (usually known as "idiots savants") who perform considerably better than the majority of intelligent people at particular intellectual tasks, including ones that involve mental arithmetic, calendar-date calculations, memorising large bodies of knowledge, or musical or artistic expertise (see, e.g., Selfe, 1977; Hill, 1978; Smith and Howe, 1985; Howe and Smith, 1988; Howe, 1989a, 1989b). In some instances mentally retarded individuals do extraordinarily well at difficult mathematical problems that demand highly abstract reasoning abilities (Sacks, 1985). Similar evidence of the independence of mental abilities is provided by the finding that some autistic children can master complex mathematical and other reasoning tasks, although the same children perform very poorly at problems in which they are required to view a situation from another person's perspective. The exact reverse of this pattern of strengths and defects characterises children with Downs Syndrome (Baron-Cohen *et al.*, 1985; Leslie, 1987).

Research into the effects of brain damage provides a third, and very substantial, body of findings in support of the view that different abilities are largely autonomous. On numerous occasions it has been found that, following brain damage, certain particular abilities are totally destroyed, whilst others, even very similar ones, remain completely unimpaired. (See, for example, Ellis and Young, 1988, for a survey of this evidence.)

Fourth, in a number of experiments it has been observed that, contrary to what would be expected if different abilities were connected to or dependent upon one another, activities which impede or interfere with performance at one mental task may have little or no effect on another task. That is so even when the two tasks are highly similar. For instance, Allport *et al.*, (1972) examined performance at two skills, sight-reading of piano music and speech-shadowing. The subjects, who were experienced sight-readers, first performed each task on its own. For the sight-reading test, an unfamiliar piano piece was placed in front of the subject, who was required to play it on the piano without rehearsal. In the shadowing test the subject heard a prose passage through headphones and was required to speak it aloud at the same time. The two tasks were then combined: the subject was

asked to continue as best she could with both tasks together. Difficult as this might seem after a very small amount of practice, subjects were able to perform the two tasks together without performance at either breaking down. Yet more surprisingly, the two tasks did not even seriously affect one another: in the dual task condition each was performed almost as well as when it was done on its own. A reasonable interpretation of this finding is that the mental processing mechanisms that are required for doing the first task are essentially separate from the ones that underly performance at the second task. Admittedly, in many cases doing one task does interfere with performance at another, but is probably because both depend upon shared mental processes, shared knowledge, or shared sub-skills.

Fifth, research into the acquisition of exceptional abilities has yielded a number of findings that are contrary to what would be expected if it were true that specific abilities are controlled or constrained by general intelligence. For example, (a) in certain circumstances ordinary people can master skills or gain intellectual abilities that are regarded as being quite exceptional, but (b) when this happens the individual's other abilities remain unaffected. For instance, Chase and Ericsson (1981) trained a young adult of normal intelligence to recall lists of digits. Eventually, after a two-year period, he could recall lists that contained about ten times as many items as the median maximum list length that can be recalled without error. However, the researchers found that the training had no effect on their subject's ability to recall other kinds of information. Both the huge training effect and the total lack of transfer appear to contradict the notion that particular intellectual skills are tied to or constrained by an individual's general ability.

Sixth, there is a growing body of evidence showing that the extent to which individuals can succeed at a variety of difficult tasks which demand extremely complex cognitive functioning is on the one hand independent of measured intelligence level, and on the other hand remarkably specific to the contexts of the tasks and the particular domains of knowledge a person happens to possess. Hence, for instance, some individuals with very low IQs who have a long-standing interest in horse-racing are capable of highly complex multiplicative reasoning involving multiple interaction effects, providing that the reasoning task is closely related to their field of interest (Ceci and Liker, 1986). Similarly, performance at certain cognitively complex tasks that people encounter at work is also unrelated to measured intelligence (Klemp and McClelland, 1986; Scribner, 1986). The same is true of complex skills in everyday life (Lave *et al.*, 1984).

In addition to the findings in the above six categories, the results of a number of studies have produced further data that appear to contradict the view that general intelligence constrains particular abilities. For instance, children who are given formally identical tasks, equivalent in the extent to which they demand complex mental functioning, perform up to six times as well when the tasks either invoke the children's existing interests or are done in the familiar context of the child's own home, than when these conditions are not met (Ceci and Bronfenbrenner, 1985; Ceci *et al.*, 1988). Various other research findings provide additional evidence concerning the autonomy of particular mental abilities (see, for example, Zigler and Seitz, 1982; Klemp and McClelland, 1986; Horn, 1986). Also, in numerous instances cross-cultural studies have demonstrated that abilities which in one culture are very rare, and regarded as necessitating quite exceptional mental ability, are found to be quite commonplace in a different culture, and in no way indicative of above-average intelligence (Laboratory of Comparative Human Cognition, 1979, 1983). This is yet another fact that is hard to rec-

oncile with the view that any kind of general ability, intelligence, or mental capacity controls or constrains specific abilities.

Taken together, the various research findings make a formidable case in support of the position that the different mental abilities of a single person are to a marked extent separate, autonomous, and independent of one another, and that they are not to any substantial extent either controlled or constrained by intelligence or "general ability". It appears that there is no compelling reason for believing in the existence of some kind of unitary mental ability or intelligence, except at a purely descriptive level.

HOW DO WE ACCOUNT FOR CORRELATIONS BETWEEN LEVELS OF ABILITY?

Convincing as the above conclusion may seem, it appears to be falsified by the fact that positive correlations (some of which are substantial) between a person's scores at different mental tasks are commonplace. There exists a wealth of correlational evidence that has been widely interpreted as ruling out the possibility of human abilities being highly specific or autonomous. Moreover, the shared variance in performance scores at different intellectual tests makes it possible to derive various algebraic factors, including the general factor ("g"). The fact that g (or a "positive manifold") can be derived is sometimes regarded as verifying the existence of general intelligence as a real entity (Eysenck, 1988). But L. L. Thurstone challenged that view in the 1930s, and the interpretation of factor-analytic findings has continued to be a matter of considerable debate. A number of researchers have argued that the existence of a g factor has little direct bearing on questions about the causes of individual differences in intellectual abilities. (See, for example, Brynner and Romney, 1986; Horn, 1986; Rabbitt, 1988a).

However, there are other kinds of evidence, all taking the form of correlations in task performance scores, that appear to contradict the notion that a person's different abilities can be largely independent. For example, first, that view appears to be negated by the simple fact that there are substantial positive correlations between any individual's scores at different mental tasks, such as the ones that form the different subtests which contribute to measures of intelligence or general ability. Second, if different abilities are largely independent, how can one explain the well-substantiated finding that measured intelligence is correlated with performance scores at a variety of simple information-processing tasks (such as those which assess the amount of time a person takes to identify letters or compare the length of lines, for instance) that purport to assess the efficiency of fundamental cognitive processes? (See Nettelbeck, 1987, for a useful review.) And third, how is it possible to explain the fact that intelligence test scores are also correlated with "evoked potential" measures, which reflect aspects of the brain's physiological functioning (Eysenck and Barrett, 1985)? Evidence of these kinds has led Eysenck (1988) to assert that the correlation between reaction time and intelligence is a reflection of whatever is common to all mental tests. He regards speed of mental processing as the fundamental variable underlying differences in general intelligence.

In fact, however, the contradictions are more apparent than real. The finding that performance levels at a number of tasks are correlated is not at all incompatible with the view that the different cognitive abilities underlying the tasks are largely independent. Correlational and factor-analytic techniques can merely establish whether or not relationships exist between task scores. They can detect structures or patterns in performances, but as a rule they cannot provide definitive evidence concerning the underlying causes.

How can the correlations be accounted for, if they do not reflect the influence of some kind of general ability or intelligence that affects performance at a variety of intellectual tasks? One contributing reason is that any two tasks often contain elements that are common to both, or depend on the same knowledge or the same skills or processes. Also, performance at each of two or more tasks may be affected by any of a variety of personal traits, qualities, and attributes. The number of different phenomena that can influence performance at any mental task is extremely large. For that reason it is not at all surprising that, when performance is tested at two different tasks, one or more of the phenomena that affect how the first task is done will also influence the same person's performance at the second task. Whenever that happens, a consequence is that there will be a correlation between an individual's scores at the two tasks.

Conceivably, the complications introduced by the fact that different tasks have common elements could be removed by making use of tasks that are extremely simple or "basic". In theory, it might be possible to devise a task which is so simple that someone's level of performance at it provides a measure of that individual's competence at a particular cognitive skill, and nothing else. If that were possible, a correlation between a person's performance levels at two such tasks might legitimately be regarded as providing evidence for the existence of the influence of some kind of general ability. In practice, however, with none of the tasks that experimenters have been able to devise has it proved possible to approach at all closely the state of affairs in which someone's performance at any task provides an uncontaminated measure of any one basic skill.

To complicate matters, it turns out that even with tasks using materials that have been deliberately chosen to be highly familiar to all subjects, in order to rule out the possibility that differences between subjects in extraneous factors such as their knowledgeability about the task elements can affect performance (and thereby confound the interpretation of the findings), this intention has proved extremely difficult to achieve. For example, even with very simple and familiar items such as digits, performance at tasks that make use of them is affected by practice (Rawlings *et al.*, 1989), is related to age (Ceci and Tishman, 1982) and is influenced by ability- and age-related differences in the way in which knowledge about those items is cognitively represented (Chi and Ceci, 1987).

Even if it did prove possible to devise very simple tasks for which there was no ambiguity about the cognitive processes involved, the problems for anyone who wished to use measures of performance at those tasks as evidence for or against the influence of some kind of general ability would not be entirely resolved. The reason is that there remains a large number of other influences, in addition to those that are strictly cognitive. These additional influences may exert similar effects on performance at a number of different tasks. If they do, there will be correlations in task scores. The existence of these other influences provides another reason why it would be wrong to assume that correlation provided unambiguous evidence of the influence of a general ability.

The other influences include a variety of states and attributes, most of which are related more to temperament, mood, or personality than to cognitive processes as such. Amongst them are, for instance, motivation, fatigue, degree of interest, attentiveness, competitiveness, "test-wiseness", the extent of a person's involvement with the task, impulsivity, co-operativeness, perseverance, and self-confidence. Other factors that can affect performance include perceptual sensitivity, adaptation effects, and influences that may reflect noise level within the perceptual system (Nettlebeck, 1987). Any of the above may have a similar influence

on an individual's performance at more than one task. Partly for that reason, the possible reasons for any correlation in measurements of performance at different tasks are numerous. Consequently, even if it were possible to devise tasks which shared no common elements whatsoever, the assumption that the cause of the correlations must lie in the shared dependence of each task on some entity of general ability or intelligence would still be highly questionable.

The objections to the use of correlational findings as evidence to support the view that there does exist some kind of general ability would be weaker if the correlations in performance at different intellectual tasks were very substantial. But they are not, except in cases where the cognitive tasks are highly complex, or heavily dependent upon previous learning. In such instances it is especially likely that the correlations in a person's levels of performance at a number of tasks are due to the presence of learned knowledge or mental skills that contribute to more than one of the tasks.

A possible counter-argument is that there have been some reported findings of large correlations between overall intelligence test scores and level of performance at very simple cognitive tasks. Typically, reaction times or inspection times, or measures of evoked potentials, are calculated for tasks that require a subject to encode, match, identify, compare, rotate, or retrieve items of information. However, in all such cases it is true either that other researchers have been unable to replicate the findings, or the subject samples have been very small and have included a disproportionately large number of mentally retarded individuals (Howe, 1988a, 1988b), or that the correlations are based upon sets of composite tests scores rather than measures of performance at single tasks. Eysenck and Barrett (1985) note that the findings are confusing, partly because "many investigators have been relatively incompetent and slapdash". Moreover, few researchers have taken into consideration the major complications that are introduced by the fact that measures of response speed do not provide reliable indications of processing speed, and by the presence of interactions between processing speed and measured intelligence (Rabbitt, 1988b). In those studies that have been free of methodological defects and have been based on representative samples that satisfactorily match the distribution of measured intelligence in the general population, the observed correlations have usually proved to be very small, typically around -0.03 or less, and often zero. (See, for example, Smith and Stanley, 1983; Irwin, 1984; Keating *et al.*, 1985; Ruchalla *et al.*, 1985; Anderson, 1986; Eysenck, 1986.)

IMPLICATIONS

If none of the objections to the view that mental abilities are considerably more independent than has been acknowledged can be sustained, where does this leave the concept of intelligence, in relation to children's education? In one important respect, it is unaffected. So far as some of the practical uses of measures of general ability are concerned, what has been said makes virtually no difference at all. Just as it is useful in everyday life to have terms such as "intelligent" and "bright" that provide a brief, generalised summary assessment of someone's intellectual abilities, for applied purposes it is often helpful to be able to draw upon tests which, by sampling intellectual skills, provide information of considerable practical value.

Current intelligence tests and ability scales which assess intellectual abilities form effective and economical devices for sampling the intellectual skills that contribute to a person's success in real-life situations encountered within educational contexts. In the field of education, the predictive value of such a test can be

considerable. Exactly how effective a particular test actually is for achieving particular aims is a straightforwardly empirical question. Broadly speaking, intelligence test scores are useful for predicting success at school and at achievements that depend on school success. Equally, intelligence test scores are strongly influenced by schooling (see Heyns, 1978; Howe, 1972; Jencks *et al.*, 1972; Jencks and Crouse, 1982; Lave, 1977, Schmidt, 1967, for example). But such scores are not nearly so useful for predicting success at tasks that are not directly related to schooling or are encountered outside educational contexts (Zigler and Seitz, 1982).

The point at which intelligence test scores become counterproductive is when they cease to be used purely *descriptively* and are introduced in ways which falsely imply that intelligence is also an *explanatory* concept. I have argued elsewhere (Howe, 1988a; see also Howe, 1976) that for the concept of intelligence to be genuinely explanatory, (in which event an indication of someone's measured intelligence would not only describe what that person could do but would also give an indication of the underlying reasons for that level of performance being achieved), at least one of ten specified conditions would have to be met, and that in fact none of those conditions is achieved.

In educational practice the borderline between legitimate descriptive uses of intelligence measures and quasi-explanatory misuses can be a narrow one. It is the difference between the teacher who says that Josephine's test scores suggest that she currently lacks intellectual skills which are necessary in order to achieve at a certain level, and the teacher who claims that Josephine's test scores prove that she has insufficient intelligence or "potential" ever to succeed. Scores obtained on a test are never more than measures of what a person can actually do at the time. Descriptive information of that kind does often form a basis for making valuable, and fairly accurate, predictions. But because the information is only descriptive, and not explanatory, it can only indicate what is *likely* in the future, given that the underlying conditions that influence at individual's current test score remain stable. It cannot set limits on what is *possible*. Contrary to what is often believed, scores at tests measuring mental abilities do not indicate limits imposed by some kind of fixed potential for future achievement.

Does it actually matter, so far as practical educational issues are concerned, if people wrongly assume that there exists a unitary and constraining entity of general intelligence which accounts for and explains differences between individuals in their intellectual ability? It is important to be aware of the fallacies underlying this view. The following are just some of the reasons.

First, once it is apparent that what a child is capable of doing largely depends upon particular abilities which that individual has acquired, rather than upon the presence or absence of some mythical quality or power of intelligence (or "intelligences", in Gardner's 1984 conceptualisation, for that matter), it becomes clear that the way to help the child to gain a valued intellectual ability is to maximise the chances of the individual acquiring the particular skills and the particular kinds of knowledge which that ability depends upon. In deciding what is required, the concept of intelligence is redundant and unnecessary. Introducing it only confuses the issue.

Second, it becomes apparent that the likelihood that training will "transfer" from one ability to another will depend upon the extent to which a child possesses knowledge, or skills, or other attributes, that contribute to both, rather than upon some ill-specified general quality of the person.

Third, once there is a realistic awareness of the actual status of the concept of intelligence it is apparent that a measure of intelligence can never, on its own, provide a valid indication of any child's "potential" or "ceiling". Such a measure simply describes what an individual does on a particular occasion, when faced with certain tasks.

Fourth, it is also clear that the degree to which a child has previously been successful in one sphere of endeavour cannot provide a reliable indication of the child's likely degree of success at a different area of activity. Whilst it is true to say that there exist certain mental skills and strategies that are broadly applicable to each of a range of tasks, an individual's capacity to deploy acquired mental skills is often highly specific to particular tasks, contexts and domains of knowledge. The fact that intelligence is not unitary, together with the finding that different skills tend to be independent, autonomous, and controlled by mental computing systems that are to some extent modular (Fodor, 1983), have formed something of a bugbear to educators who have designed programmes that have the apparently reasonable aim of teaching broadly applicable general-purpose thinking skills and mental strategies (Glaser, 1984). We shall have to live with the fact that the domain-specificity of human abilities restricts the range of applicability of mental skills.

Fifth, once it is apparent that intelligence is a solely descriptive concept it also becomes clear that there is little point in judging the effectiveness of any educational programme or intervention by assessing its effects on intelligence test scores. For educators to do so is rather like factory managers putting their main efforts into making sure that their factory gains the highest possible score on a test that has been designed to provide a convenient and inexpensive estimate of the factory's productivity, rather than directing their energies towards ensuring that the factory actually produces more goods.

In the present author's view, the belief that intelligence exists at all, except as a solely descriptive term, is nothing but a figment of twentieth-century psychologists' imaginations. Its existence, like that of a number of concepts that seemed real to previous generations of scientists, including "phlogiston", the four "humours", and "the ether", may soon be seen as having been illusory. So far as attempts at scientific explanation are concerned, intelligence, like those other terms, may quickly become as dead as the dodo, and belong only to history.

Correspondence and requests for reprints should be addressed to Dr M. J. A. Howe, Department of Psychology, University of Exeter, Washington Singer Laboratories, Exeter, EX4 4QG.

REFERENCES

- ALLPORT, D. A., ANTONIS, B., and REYNOLDS, P. (1972). On the division of attention: a disproof of the single channel hypothesis. *Q. Jl. exp. Psychol.*, 24, 225-235.
- ANDERSON, M. (1986). Inspection time and IQ in young children. *Person. individ. Diff.*, 7, 677-686.
- BARON-COHEN, S., LESLIE, A. M., and FRITH, U. (1985). Does the autistic child have a "theory of mind"? *Cognition*, 21, 27-46.
- BYNNER, J. M., and ROMNEY, D. M. (1986). Intelligence, fact or artefact: alternative structures for cognitive abilities. *Br. J. educ. Psychol.*, 56, 13-23.
- CECI, S. J., and TISHMAN, J. (1982). Developmental changes in encoding as a function of type of stimuli. Unpublished manuscript. Ithaca, New York: Cornell University.
- CECI, S. J., and BRONFENBRENNER, U. (1985). Don't forget to take the cupcakes out of the oven: strategic time-monitoring, prospective memory, and context. *Child Dev.*, 56, 175-190.
- CECI, S. J., and LIKER (1986). A day at the races: a study of IQ, expertise, and cognitive complexity. *J. exp. Psychol.: General*, 115, 255-266.

- CECI, S. J., BAKER, J. G., and BRONFENBRENNER, U. (1988). Prospective remembering, temporal calibration, and context. In GRUNEBERG, M. M., MORRIS, P. E., and SYKES, R. (Eds.), *Practical Aspects of Memory: Current Research and Issues*. Chichester: Wiley.
- CHASE, W. G. and ERICSSON, K. A. (1981). Skilled memory. In ANDERSON, J. R. (Ed.), *Cognitive Skills and their Acquisition*. Hillsdale, NJ: Erlbaum.
- CHI, M. T. H., and CECI, S. J. (1987). Content knowledge: its role, representation, and restructuring in memory development. *Advances in Child Dev.*, 20, 91-142.
- ELLIS, A. W., and YOUNG, A. W. (1988). *Human Cognitive Neuropsychology*. Hove, Sussex: Erlbaum.
- EYSENCK, H. J. (1986). The theory of intelligence and the psychophysiology of cognition. In STERNBERG, R. J. (Ed.), *Advances in the Psychology of Human Cognition, Volume 3*. Hillsdale, NJ: Erlbaum.
- EYSENCK, H. J. (1988). Editorial: the concept of "intelligence": useful or useless. *Intelligence*, 12, 1-16.
- EYSENCK, H. J., and BARRETT, P. T. (1985). Psychophysiology and the measurement of intelligence. In REYNOLDS, C. R., and WILLSON, V. L. (Eds.), *Methodological and Statistical Advances in the Study of Individual Differences*. New York: Plenum.
- FODOR, J. A. (1983). *The Modularity of Mind*. Cambridge, Mass.: MIT Press.
- GARDNER, H. (1984). *Frames of Mind*. London: Heinemann.
- GARDNER, H. (1988). Creative lives and creative works: a synthetic scientific approach. In STERNBERG, R. J. (Ed.), *The Nature of Creativity: Contemporary Psychological Perspectives*. Cambridge: Cambridge University Press.
- GESCHWIND, N. (1983). The organization of the living brain. In MILLER, J. (Ed.), *States of Mind: Conversations with Psychological Investigators*. London: BBC.
- GLASER, R. (1984). Education and thinking: the role of knowledge. *Am. Psychol.* 39, 93-104.
- HEYNS, B. (1978). *Summer Learning and the Effects of Schooling*. New York: Academic Press.
- HILL, A. L. (1978). Savants: mentally retarded individuals with special skills. In ELLIS, N. R. (Ed.), *International Review of Research in Mental Retardation, Volume 9*, New York: Academic Press.
- HORN, J. (1986). Intellectual ability concepts. In STERNBERG, R. J. (Ed.), *Advances in the Psychology of Human Intelligence, Volume Three*. Hillsdale, NJ: Erlbaum.
- HOWE, M. J. A. (1972). *Understanding School Learning: A New Look at Educational Psychology*. New York: Harper and Row.
- HOWE, M. J. A. (1976). Good learners and poor learners. *Bull. Br. psychol. Soc.*, 29, 16-19.
- HOWE, M. J. A. (1982). Biographical evidence and the development of outstanding individuals. *Am. Psychol.*, 37, 1071-1081.
- HOWE, M. J. A. (1988a). Intelligence as an explanation. *Br. J. Psychol.*, 79, 349-360.
- HOWE, M. J. A. (1988b). The hazards of using correlational evidence as a means of identifying the causes of individual ability differences: a rejoinder to Sternberg and a reply to Miles. *Br. J. Psychol.*, 79, 539-545.
- HOWE, M. J. A. (1989a). *Fragments of Genius: Idiots Savants and the Psychological Investigation of Remarkable Feats by Mentally Retarded Individuals*. London: Routledge.
- HOWE, M. J. A. (1989b). The strange achievements of idiots savants. In COLMAN, A. M. and BEAUMONT, J. G. (Eds.), *Psychology Survey 7*. Leicester: The British Psychological Society.
- HOWE, M. J. A., and SMITH, J. (1988). Calendar calculating in "idiots savants": how do they do it? *Br. J. Psychol.*, 79, 371-386.
- IRWIN, R. J. (1984). Inspection time and its relation to intelligence. *Intelligence*, 8, 47-65.
- JENCKS, C., and CROUSE, J. (1982). Aptitude versus achievement: should we replace the SAT? In SCHRADER, W. B. (Ed.), *New Directions for Testing and Measurement, Guidance and Program Improvement, No. 13*. San Francisco: Jossey Bass.
- JENCKS, C., SMITH, M., ACLAND, H., BANE, M. J., COHEN, D., GINTIS, H., HEYNS, B., and MIDTCHELSON, S. (1972). *Inequality: A Reassessment of the Effects of Family and Schooling in America*. New York: Basic Books.
- KEATING, D. P., LIST, J. A., and MERRIMAN, W. E. (1985). Cognitive processing and cognitive ability: a multivariate validity investigation. *Intelligence*, 9, 149-170.
- KLEMP, G. O., and MCCLELLAND, D. C. (1986). What characterizes intelligent functioning among senior managers? In STERNBERG, R. J. and WAGNER, R. K. (Eds.), *Practical Intelligence: Nature and Origins of Competence in the Everyday World*. Cambridge: Cambridge University Press.
- LABORATORY OF COMPARATIVE HUMAN COGNITION (1979). What's cultural about cross-cultural psychology? *Ann. Rev. Psychol.*, 30, 145-172.
- LABORATORY OF COMPARATIVE HUMAN COGNITION (1983). Culture and cognitive development. In MUSSEN, P. H. (Ed.), *Handbook of Child Psychology, Fourth Edition, Volume One*. New York: Wiley.
- LAVE, J. (1977). Tailor-made experiments and evaluating the intellectual consequences of apprenticeship training. *Quarterly Newsletter of the Institute for Comparative Human Development*, 1, 1-3.
- LAVE, J., MURTAUGH, M., and DE LA ROCHE, D. (1984). The dialectic of arithmetic in grocery shopping. In ROGOFF, B. and LAVE, J. (Eds.), *Everyday Cognition: Its Development in Social Contexts*. Cambridge, Mass.: Harvard University Press.
- LESLIE, A. M. (1987). Pretense and representation: the origins of "theory of mind". *Psychol. Rev.*, 94, 412-426.

- NETTLEBECK, T. (1987). Inspection time and intelligence. In VERNON, P. A. (Ed.), *Speed of Information Processing and Intelligence*. Norwood, NJ.: Ablex.
- RABBITT, P. M. A. (1988a). Critical notice of R. J. Sternberg (Ed.): *Handbook of Human Intelligence: Advances in the Psychology of Human Intelligence*, Vols. 1 and 2; *Human Abilities; Beyond IQ: A Triarchic Theory of Human Intelligence*. *Q. Jl. exp. Psychol.*, 40A, 167-185.
- RABBITT, P. M. A. (1988b). Does fast last? Is speed a basic factor determining individual differences in memory? In GRUNEBERG, M. M., MORRIS, P. E., and SYKES, R. (Eds.), *Practical Aspects of Memory: Current Research and Issues*. Chichester: Wiley.
- RAWLINGS, P., WALKER, E., and CECI, S. J. (1989). Semantic priming effects in backward masking: implications for *g*. *Intelligence*, 13.
- ROZIN, P. (1976). The evolution of intelligence and access to the cognitive unconscious. *Progress in Psychobiology, Physiology and Psychology*, 6, 245-280.
- RUCHALLA, E., SCHAT, E., and VOGEL, F. (1985). Relations between mental performance and reaction time: new aspects of an old problem. *Intelligence*, 9, 189-205.
- SACKS, O. (1985). *The Man Who Mistook His Wife for a Hat*. London: Duckworth.
- SCHMDT, W. H. O. (1967). Socio-economic status, schooling, intelligence, and scholastic progress in a community in which education is not yet compulsory. *Paedog. Eur.*, 2, 275-286.
- SCRIBNER, S. (1986). Thinking in action: some characteristics of practical thought. In STERNBERG, R. J. and WAGNER, R. K. (Eds.), *Practical Intelligence: Nature and Origins of Competence in the Everyday World*. Cambridge: Cambridge University Press.
- SELFE, L. (1977). *Nadia: A Case of Extraordinary Drawing Ability in an Autistic Child*. London: Academic Press.
- SMITH, G. A., and STANLEY, G. (1983). Clocking *g*: relating intelligence and measures of timed performance. *Intelligence*, 7, 353-368.
- SMITH, J., and HOWE, M. J. A. (1985). An investigation of calendar-calculating skills in an "idiot savant". *Int. J. Rehab. Res.*, 8, 77-79.
- ZIGLER, E., and SEITZ, V. (1982). Social policy and intelligence. In STERNBERG, R. J. (Ed), *Handbook of Human Intelligence*. New York: Cambridge University Press.

(Manuscript received 17th October, 1988)

RESEARCH NOTES

Br. J. educ. Psychol., 59, 361-365, 1989

ONE IN ELEVEN: PUPILS REQUIRING SPECIAL EDUCATIONAL ASSISTANCE IN CATHOLIC SCHOOLS IN THE AUSTRALIAN STATE OF VICTORIA

By CHRISTOPHER SZADAY, DES PICKERING AND PAUL DUERDOTH
(*Faculty of Special Education and Paramedical Studies, Victoria College — Burwood Campus, Melbourne, Australia*)

SUMMARY. The perceptions of 4,353 teachers in 463 primary and secondary Catholic schools in the Australian state of Victoria about the special educational needs and additional educational requirements of 120,344 pupils are reported in this paper. One in eleven pupils in regular schools was perceived by teachers to be experiencing educational difficulties associated with traditional disability categories sufficient to require forms of additional educational assistance in order to more fully participate and succeed in the regular school programme.

INTRODUCTION

An advertising agency could not have achieved greater success than the Warnock Report (1978) in establishing in a target audience the belief that "one in six children . . . will require some form of special educational provision" (p. 41). Titles of recent books, such as *Warnock's Eighteen Per Cent* (Gipps *et al.*, 1987), and chapters in books, "Warnock's 20 per cent" (Galloway, 1985), reflect the importance of this estimate of the prevalence of pupils with special educational needs. It is not surprising that this estimate has considerable influence in Australia, even though it is based on British research findings.

In the context of the resource and staffing implications of such an estimate, there has been in Australia an increasing questioning of its validity. A recent meeting of Australian state and federal Ministers of Education is reported to have recommended that this estimate "play no role in national planning for student support services or for planning and delivery by individual systems" (Gow *et al.*, 1987, p. 36).

Another government report in Australia (Collins, 1984) has expressed dissatisfaction with the concept of "special educational need", favouring instead an elaboration of the "additional educational requirements" of pupils to increase participation and success in the regular school programme. However, as noted in this report, there is "no adequate data base for identifying . . . the levels of additional educational service requirements necessary to integrate children . . . into the regular school setting" (Collins, 1984, p. 91). Jenkinson (1988) notes that few pieces of special education research have addressed the question of the forms of special educational provision required by pupils with special educational needs.

The present study is an attempt to explore the relationship between the special educational needs and additional educational requirements of pupils in an Australian school system. Teachers were surveyed to determine: (a) the prevalence of pupils in regular schools experiencing educational difficulties associated with traditional categories of impairment, disability or handicap; and (b) the forms of additional educational assistance required by such pupils to more fully participate and succeed in the regular school programme.

METHOD

As part of a larger study into the special educational needs of Catholic schools in the Australian state of Victoria (Pickering *et al.*, 1988), teacher questionnaires were distributed in April, 1987, to all 495 Catholic primary and secondary schools. The distribution of these schools on indices of socio-economic status and ethnic origin of pupils closely approximates that of all Australian schools. School principals were encouraged by the Catholic Education Office of Victoria to return completed questionnaires to the authors. These

questionnaires were to be completed on all pupils from preparatory grade to Year 10, the years of compulsory schooling.

Completed questionnaires were received from 4,353 primary class teachers and secondary form teachers in 463 schools. A higher response rate was received from primary schools (98.7 per cent), with whom the Catholic Education Office of Victoria had more influence, than secondary schools (75.9 per cent). Teacher ratings were thus provided on 120,344 pupils, constituting 76.7 per cent of the eligible pupil population of Catholic schools in Victoria.

Teachers were requested to list each pupil in their class or form group and indicate: (a) whether each pupil experienced educational difficulties associated with one or more of the traditional disability categories; and (b) which forms of additional educational assistance, if any, were required to enable such pupils to more fully participate and succeed in the school programme. The teacher questionnaire included a modified version of the grid suggested by Warnock (1978) for use as a basis for statistical returns. Pilot testing of this grid in November, 1986, with teachers in two primary and two secondary schools resulted in the initial ten disability categories being reduced to eight, and the provision of brief descriptions of these eight forms of disability. Warnock's (1978) five-point scale of degree of impairment was not used in this study because the focus was on the additional educational requirements of pupils with special educational needs rather than registering degree of impairment. The categories of special educational assistance included in the teacher questionnaire were adaptations and extensions of those of Stone (1984), currently used in Victorian Ministry of Education integration policy documents, and were intended to be inclusive of special education service delivery patterns to pupils in regular classrooms and schools.

The following definitions, corresponding to the traditional categories of disability, were provided to teachers:

Vision Problem, where the pupil is identified as having or suspected to have a visual impairment sufficient to impede educational progress, thereby requiring supplementary means of instruction (e.g. braille, magnifying aids);

Hearing Problem, where the pupil is identified as having or suspected to have a hearing loss sufficient to impede educational progress, thereby requiring hearing aids or supplementary means of instruction;

Co-ordination Problem, encompassing pupils with severe disabilities (e.g., cerebral palsy, spina bifida) through to pupils with lesser co-ordination problems (e.g., reflected in hand-writing, athletic performance or general clumsiness) who nevertheless require extra educational assistance as a result of such problems;

Health Problem, encompassing pupils with chronic conditions (e.g., asthma, cystic fibrosis, juvenile rheumatoid arthritis, cancer) which affect educational progress because of the actual symptoms of the disease or the pupil's absence from school;

Speech and Communication Problem, of sufficient severity to impede the pupil's ability to interact and communicate with others, and which may produce negative feelings in either the pupil or listener;

General Learning Problem, where the pupil has significant difficulties across all or most of the curriculum areas, and whose achievement levels appear to reflect general ability.

Specific Learning Problem, where the pupil has a significant difficulty in spoken or written language, or reading or mathematics, which cannot primarily be attributed to a lack of general ability, sensory impairment or environmental circumstances; and

Emotional or Behaviour Problem, where the pupil's actions at school cause distress to either self, peers or teachers.

Two other categories, not traditionally the domain of special education, were included in the teacher questionnaire to distinguish those pupils whose difficulties were primarily the result of a lack of familiarity with English or who manifested a teacher-perceived motivational deficit:

English as a Second Language, where the pupil was born in a country whose major language is not English, or was born in Australia with one or both parents born overseas in a country whose major language is not English; and

Lacks Motivation or Regularly Fails to Complete Work, where the pupil appears uninterested in school activities, may be described by teachers as "lazy", and has to be "pushed" to complete schoolwork or assignments.

Pupils nominated in either of these categories who were not nominated in any of the traditional special education categories listed above were not included in the data reported in this paper.

The forms of additional educational assistance likely to be required to increase the participation and success in the regular school programme of pupils nominated by teachers to be experiencing difficulties included on the teacher questionnaire were:

Programme Advice, where the teacher receives advice from consultants and others about ways in which the pupil's needs can best be met. This advice may be given in the form of specific in-service education programmes or may involve direct advice about a particular pupil given to teachers individually or in groups;

Curriculum Support, where particular curriculum materials, equipment or worksheets are provided to the teacher for use in the classroom;

Team-Teaching, in which another adult (specialist teacher or aide) works in the classroom alongside the class teacher, assisting in a variety of classroom activities, but is not specifically "tagged" to a particular pupil;

In-Class Support, where a special teacher or aide works in the classroom assisting ("tagged to") a particular pupil with activities specific to his or her needs;

Withdrawal of the pupil into a room other than the normal classroom for intensive instruction, testing or therapy, either individually or in a group.

Personal Counselling of pupil as a result of home-based or school-related adjustment difficulties, by either members of the school staff or outside agencies; and

Vocational Guidance by careers teacher or other specialist to increase the pupil's awareness of career options or vocational skill.

RESULTS

Teachers nominated 13.8 per cent of the 120,344 pupils considered in this study to be experiencing educational difficulties associated with one or more of the traditional categories of disability. Primary school teachers nominated a higher percentage of pupils (14.7 per cent) than secondary school teachers (11.5 per cent). However, specific forms of additional educational assistance were nominated by teachers for only some of these pupils. Forms of additional educational assistance were specified for 9.5 per cent of pupils (approximately one in eleven), with a greater proportion of primary school pupils (10.2 per cent) than secondary school pupils (7.7 per cent) being nominated. Table 1 presents the percentage of pupils in primary and secondary schools nominated by teachers to require each of the forms of additional educational assistance in each of the disability categories. As pupils could be nominated by teachers in more than one category of disability and be nominated as requiring more than one form of additional educational assistance, the sum of the cells in Table 1 exceeds the total number of students so designated.

More pupils were considered by teachers to experience learning or behaviour problems than sensory, physical or language disorders, although there was a marked overlap between categories. Almost one third (31.6 per cent) of the pupils were indicated by teachers in more than one disability category.

Responding teachers nominated the seven forms of additional educational assistance in similar proportions for each of the disability categories. Primary school teachers requested in-class support as the major form of assistance in seven of the eight disability categories, while secondary school teachers nominated this form of assistance in only two categories. Secondary school teachers were more likely to choose withdrawal as the major

TABLE 1

PERCENTAGE OF PRIMARY PUPILS (N = 87,670) AND SECONDARY PUPILS (N = 32,674) REQUIRING FORMS OF ADDITIONAL EDUCATIONAL ASSISTANCE BY CATEGORY OF DISABILITY AND LEVEL OF SCHOOLING

Category	School	N	Total	Form of additional educational assistance						
				PA	CS	TT	IS	WI	PC	VG
Vision	Pr	502	0.57	0.19	0.21	0.20	0.29	0.24	0.12	0.03
	Sec	118	0.36	0.07	0.11	0.11	0.13	0.17	0.09	0.04
Hearing	Pr	447	0.51	0.18	0.15	0.18	0.24	0.20	0.09	0.02
	Sec	64	0.20	0.06	0.07	0.08	0.11	0.09	0.06	0.03
Coordination	Pr	1,350	1.53	0.48	0.54	0.54	0.78	0.66	0.31	0.07
	Sec	174	0.55	0.13	0.16	0.19	0.24	0.29	0.17	0.10
Health	Pr	511	0.58	0.18	0.18	0.20	0.27	0.18	0.16	0.02
	Sec	220	0.68	0.10	0.12	0.25	0.17	0.19	0.17	0.12
Speech	Pr	1,547	1.76	0.56	0.52	0.59	0.87	0.80	0.32	0.06
	Sec	169	0.52	0.11	0.17	0.18	0.27	0.25	0.14	0.10
General learning	Pr	4,282	4.90	1.29	1.38	1.62	2.51	2.20	0.70	0.13
	Sec	1,328	4.01	0.63	0.89	1.40	1.73	2.11	0.83	0.56
Specific learning	Pr	2,009	2.28	0.69	0.74	0.78	1.12	1.08	0.32	0.08
	Sec	615	1.90	0.27	0.50	0.70	0.81	0.98	0.30	0.20
Behaviour	Pr	2,707	3.09	0.83	0.60	0.84	1.21	0.90	1.54	0.12
	Sec	748	2.29	0.26	0.28	0.59	0.60	0.68	1.63	0.30
Total pupils	Pr	8,905	10.16	2.58	2.48	3.07	4.34	3.72	1.98	0.22
	Sec	2,507	7.67	1.07	1.51	2.48	2.70	3.17	2.18	0.92

Note: PA = Programme Advice. CS = Curriculum Support. TT = Team-Teaching.
 IS = Class Support. WI = Withdrawal. PC = Personal Counselling.
 VG = Vocational Guidance.
 Modal frequencies are printed in bold type.

form of required assistance, and did so in four of the categories. Primary and secondary school teachers nominated personal counselling as a preferred form of assistance for pupils designated as manifesting emotional or behaviour problems. Two or more forms of additional assistance were nominated for approximately half (47.4 per cent) of the pupils experiencing difficulty.

DISCUSSION

The research methodology described in this paper was an attempt within an epidemiological research framework to bridge the ideological, theoretical and professional gap between the traditional practice of identifying special educational needs, usually through specification of disability (Warnock, 1978), and the more radical suggestion of the specification of additional educational requirements within a non-categorical approach to special educational provision (Collins, 1984).

Teachers differentiated between pupils who experience educational difficulties associated with one or more of the traditional disability categories (visual impairment, hearing impairment, physical disability, health impairment, language disorder, intellectual disability, learning disability and emotional disturbance) whose needs can be met by classroom teachers alone and those pupils for whom special educational assistance is required. This study suggests that some prevalence estimates of those pupils designated as having special educational needs may be confusing these two groups of pupils.

One pupil in eleven was perceived by teachers in Catholic schools in the Australian state of Victoria (one in ten in primary schools and one in 13 in secondary schools) to require additional educational assistance in order to participate more fully in the regular

school programme as a result of an impairment, disability or handicap. However, this estimate must take into account the lower response rate of secondary schools in this study and that the nature of the additional educational assistance required by secondary pupils may vary between subjects and teachers.

Teachers nominated a variety of forms of additional educational assistance to increase the participation and success in the regular school programme of such pupils. The most frequently requested services involve the direct participation of special education service providers in the environment in which the problems are manifested — the classroom — or direct forms of assistance to the pupils experiencing difficulty (e.g., withdrawal and counselling). At a time when greater emphasis is being placed by special educators and educational psychologists on consultancy and in-service training activities, the results of this study of teacher perceptions suggest alternative patterns of special education service delivery to regular classrooms and schools.

Correspondence and requests for reprints should be addressed to Christopher Szaday, Faculty of Special Education and Paramedical Studies, Victoria College, Burwood Campus, 221 Burwood Highway, Burwood, Victoria 3125, Australia.

REFERENCES

- COLLINS, M. K. (Chair) (1984). *Integration in Victorian Education*. Report of the Ministerial Review of Educational Services for the Disabled. Melbourne: Government Printer.
- GALLOWAY, D. (1985). *Schools, Pupils and Special Educational Needs*. London: Croom Helm.
- GIPPS, C., GROSS, H., and GOLDSTEIN, H. (1987). *Warnock's Eighteen Per Cent: Children with Special Needs in Primary Schools*. London: Falmer Press.
- GOW, L., SNOW, D., BALLA, J. and HALL, J. (1987). *Report to the Commonwealth Schools Commission on Integration in Australia (Volume 1)*. The University of Wollongong.
- JENKINSON, J. C. (1988). *School and Disability: Research and Practice in Integration*. Hawthorn: Australian Council for Educational Research.
- PICKERING, D., SZADAY, C., and DUERDOTH, P. (1988). *One in Eleven: Special Educational Needs of Catholic Schools in Victoria*. Melbourne: Victoria College.
- STONE, K. (1984). Watching them grow . . . Cobram Special Developmental School: supporting the integration process. *Inter View*, No. 14, pp. 4-9.
- WARNOCK, H. M. (Chairman) (1978). *Special Educational Needs*. Report of the Committee of Enquiry into the Education of Handicapped Children and Young People. London: HMSO.

(Manuscript received 19th September, 1988)

Br. J. educ. Psychol., 59, 366-371, 1989

CROSS-VALIDATION OF SHORT FORMS OF THE WISC-R IN TWO BRITISH SAMPLES

By JANET HUNTER*, WILLIAM YULE, MARIE ANNE URBANOWICZ

AND RICHARD LANSDOWN**

(*University of London, Institute of Psychiatry*)

SUMMARY. Short forms of the WISC-R were computed by multiple linear regression on two samples of British children aged 6 to 12 years. The predictive validity of each short form was assessed both within the sample on which it was developed and cross-validated on the other sample. The empirically determined loss of power of prediction was found to be less than expected on a priori grounds. Predictions of individual scores at the lower end of the IQ distribution were less acceptable. It is concluded that short forms are robust and of value in research and for screening purposes, but cannot be recommended for clinical purposes.

INTRODUCTION

The Wechsler Intelligence Scales for Children —Révised are still very widely used throughout the world for both research and clinical purposes. This is largely because of the vast data base which has accumulated on their use. However, they are time consuming — and hence expensive — to use, taking between one and one-and-a-half hours to administer all the subscales. For this reason, there have been many attempts to develop shortened forms of the scales to be used at least for screening purposes (Silverstein, 1975, 1982; Kaufman, 1979; Quatrocchi and Sherrets, 1980; Beck *et al.*, 1983).

It is generally recognised that shortened forms of any test will be less reliable and less valid than the full version. What is often overlooked is that whatever validation studies are undertaken, their results may not be generalisable to specific subsamples of the population. Typically, short forms are developed on the basis of the original standardisation data in the hope that they can be applied to children referred to various clinics. As Phillips (1984) recently reminded us, the factor structure of the WISC-R scores of clinic referrals differs from that of children in the standardisation sample. In Phillips' sample (studying data from the original WISC), a two subtest short form (Block Design and Object Assembly) gave as good estimates of Performance IQ in a general sample as in his clinic sample, but the short form of the verbal scale (Similarities and Vocabulary) did not work equally well in both samples. Not only were Verbal IQs of the clinic sample overestimated, the interpretation of Verbal-Performance discrepancies in the clinic sample was called into question.

Beck *et al.* (1983) draw attention to another aspect of cross-validation studies. Most short forms are developed on either the published standardisation data or on relatively large samples of protocols obtained for other purposes. The shortened version is developed by examining the correlation of subscale scores with Full Scale IQ and then selecting varying numbers of subtests. The sum of the subscale scores is then prorated (what Silverstein, 1984, calls using a linear scaling procedure) or are entered into a multiple-regression equation to obtain an estimated Full Scale IQ. Typically, such estimates correlate over 0.90 with the actual Full Scale IQs. However, this method is merely correlating part of the test with the whole test and capitalises on correlated errors of measurement (Kaufman, 1977). As Beck *et al.* (1983) argue, what is needed is to develop a short form on one sample and then to test out the validity on a separate one. In their own study, they did this on two samples of 300 children and found that four and five subtest short forms correlated in excess of 0.95, yielded FSIQ means which scarcely differed from those actually measured, and standard errors which were only marginally above those published for the entire WISC-R.

*Inner London Education Authority Statistics Department

**Hospital for Sick Children, Great Ormond Street, London

To date, the Beck *et al.* (1983) paper is the sole example in the literature reporting data on this cross-validation issue. The present paper addresses this issue with data from two British samples.

METHOD

As part of our studies of the effects of lead on children's development (Yule *et al.*, 1981; Hunter *et al.*, 1985; Lansdown *et al.*, 1986) WISC-R data were gathered on 194 children aged 6 to 12 years in an outer London Borough and 302 children also aged 6 to 12 years in Leeds. In the first case, the children were selected for study because they lived near a busy road; in the second sample, they were selected because of higher risk of exposure to industrial emissions of lead. In neither sample were the lead levels elevated above the EEC reference levels. The London children came from predominantly skilled working class homes; the Leeds sample had a higher proportion (23 per cent) of Registrar General Social Class IV and V families. The mean Full Scale IQ of each sample was 105.24 and 100.20 respectively. Testing was carried out at school by qualified clinical and educational psychologists who administered 11 subtests of the WISC-R (i.e., omitting Mazes), following standard procedures. All protocols were independently checked before the data were analysed.

RESULTS

Following the approach adopted by Beck *et al.* (1983), a stepwise regression procedure was used to develop short forms of the WISC-R using four or five subtest scores with FSIQ as the criterion variable. In order to assess empirically the degree of shrinkage (or loss of power) suffered by a predictive equation derived from one sample and applied to another, the regression formula from each sample was used to compute a set of predicted FSIQs for the other, and Pearson correlations between these scores and the measured FSIQs of the sample were computed, together with the standard errors of the estimates.

Table 1 shows the regression equations used to predict FSIQ in our samples. Table 2 details the Pearson correlations between the predicted and observed FSIQs using all the different prediction equations on both samples.

TABLE 1

SHORT-WISC: REGRESSIONS OF 4 OR 5 SUBSCALE SCORES ON FSIQ

<i>Derivation</i>	
London II (4 var.):	$1.60 \text{ X INFOR} + 1.60 \text{ X OBJASS} + 1.49 \text{ X ARITH} + 1.47 \text{ X COMPRE} + 39.65$ (86% variance)
London II (5 var.):	$1.36 \text{ X INFOR} + 1.47 \text{ X OBJASS} + 1.44 \text{ X ARITH} + 1.32 \text{ X COMPRE} + 1.05$ $\text{X PICARR} + 34.04$ (91% variance)
Leeds (4 var.):	$1.60 \text{ X COMPRE} + 1.60 \text{ X BLDES} + 1.58 \text{ X INFOR} + 1.37 \text{ X PIC-COM} +$ 38.29 (88% variance)
Leeds (5 var.):	$1.51 \text{ X COMPRE} + 1.43 \text{ X BLDES} + 1.52 \text{ X INFOR} + 1.26 \text{ X PICCOM} + 0.91$ $\text{X CODING} + 33.83$
Beck <i>et al.</i> 1983	
Sample 1: 4 (var.)	$2.22 \text{ X VOCAB} + 1.50 \text{ X BLDES} + 1.34 \text{ X PICARR} + 0.90 \text{ X CODING} +$ 38.98
Beck <i>et al.</i> 1983	
Sample 2: (4 var.)	$2.03 \text{ X VOCAB} + 1.45 \text{ X OBJASS} + 1.15 \text{ X PICARR} + 1.19 \text{ X ARITH} + 38.35$

TABLE 2
PERFORMANCE OF SHORT-WISC REGRESSION EQUATIONS ON TWO SAMPLES

Equation	Second London Sample		Leeds Sample	
	Pearson <i>r</i>	SE	Pearson <i>r</i>	SE
London II (4 var.)	0.94	5.00	0.92	5.19
London II (5 var.)	0.95	4.37	0.94	4.56
Leeds (4 var.)	0.92	5.42	0.93	5.00
Leeds (5 var.)	0.94	4.95	0.95	4.38
Beck <i>et al.</i> 1983 equat. 1 (4 var.)	0.93	5.32	0.89	6.08
Beck <i>et al.</i> 1983 equat. 2 (4 var.)	0.91	5.95	0.91	5.57

All the correlations are highly significant — ranging from 0.89 to 0.95 — though they tend to be lower than those obtained by Beck *et al.* (their Table 1, p. 866). The standard errors of the estimates are slightly higher (ranging from 4.4 to 5.4) than those reported by Beck, *et al.* (1983), whose SE estimate was 4.8 for the four subtest short version and 4.1 for the five subtest version. The published SE for the entire WISC-R is 3.2.

As in Beck's study, it is clear from our results that the inclusion of additional subtests in the equations results in less shrinkage on cross-validation. In fact, the five-variable equations performed better on cross-validation (i.e., when applied to the other sample) than the four-variable equations did when applied to the sample they were derived from. However, depending on the reasons for wanting to assess FSIQ using a short form of the WISC-R, the improvement in performance of a five-variable prediction equation over a four-variable version may not be justified by the extra testing involved. It should also be noted that due to high intercorrelations between the subtests improvements in validity with addition of each extra subtest above five will become progressively smaller.

Comparisons between actual mean IQ scores and predicted mean IQ scores are shown in Table 3. The means are all very close, indicating that none of the equations produced gross over- or under-estimates of the average Full Scale IQ of the samples. However, it is also important to consider whether the predictions distort the estimates of FSIQ of children at the extremes of distribution.

To look at the effect of regression towards the mean we compared groups of children whose predicted FSIQs using the four-variable prediction equations differed by 10 points or more. As expected, overall, the children whose short-WISC IQ was over-estimated were less intelligent than the average for the group, while those whose IQ was under-estimated by at least 10 points were more intelligent than the average. However, it was also clear that the prediction equations derived from the different samples performed differently in this respect. The equation from the second London sample, with its relatively higher mean FSIQ, misclassified relatively more — (15 vs 12) of the Leeds sample's lower IQ children as 10 points or more higher, while the equation derived from the Leeds sample tended to classify more of the second London children as less intelligent by 10 points or more (17 vs 3). This observation underlines the importance of taking into account the type of population from which a short form of the WISC-R is derived before applying it, particularly if the sample under study is in any way a special population. In fact, Beck's prediction equations performed fairly well on our sample. Though the standard errors of estimates for the Leeds predictions were somewhat higher, the performance of the predictors on the second London sample was surprisingly good.

TABLE 3

COMPARISONS BETWEEN ACTUAL AND PREDICTED MEANS AND STANDARD DEVIATIONS ON FSIQ SCORES

Source of Reg. equat.	Second London Sample		Leeds Sample	
	Mean	SD	Mean	SD
Actual means	105.24	14.20	100.20	13.57
London II (4)	105.27	13.29	101.56	13.17
Leeds (4)	103.87	13.48	100.21	12.63
London II (5)	105.21	13.48	101.34	13.77
Leeds (5)	103.64	13.74	100.15	12.85
Beck <i>et al.</i> 1 (4)	105.40	12.56	100.75	12.06
Beck <i>et al.</i> 2 (4)	104.16	12.28	98.89	12.18

To get an idea of the stability of extreme scores, we looked at those children who fell into the bottom 10 per cent of the FSIQ distribution who were no longer classified in this group by the regression equations (Table 4). Overall, about one in four of the children in the bottom 10 per cent according to their measured FSIQ would not have been so identified had a short form of the test been used.

TABLE 4

PER CENT OF CHILDREN ORIGINALLY IN LOWER 10 PER CENT OF SAMPLE ON FSIQ, MISCLASSIFIED BY SHORT-WISC

Equation	Second London Sample		Leeds Sample	
	%	Ratio	%	Ratio
2nd London sample (4 var.)	22.2	4/18	20.8	5/30
Leeds sample (4 var.)	33.0	6/18	26.7	8/30
Beck <i>et al.</i> no. 1	50.0	9/18	40.0	12/30
Beck <i>et al.</i> no. 2	38.9	7/18	16.7	5/30
2nd London sample (5 var.)	27.8	5/18	16.7	5/30
Leeds sample (5 var.)	27.8	5/18	26.7	8/30

DISCUSSION

As with the Beck *et al.* (1983) study, the present study yielded a number of four and five subtest forms of WISC-R which have good predictive validity. The short forms correlated only slightly less well with FSIQs obtained on the cross-validation samples than they did in their own samples. In all cases, the correlation coefficient in the British samples exceeded 0.92. The standard errors were comparable with those reported by Beck *et al.* However, their equation did not perform as well on the present samples. In part, this could be due to the characteristics of the American samples, in that their means were at least 10 points lower than the British means and they had a more restricted range. This underlines

the importance of ensuring that prediction equations are not applied to scores of children outside the characteristics of the definitive sample.

These results support the conclusion of Beck *et al.* (1983) that short forms based on multiple regression techniques are fairly robust. These short forms perform reasonably well in samples other than those on which they are based. There is no evidence for major loss of power or "shrinkage". However, this conclusion applies to the whole sample. At the lower end of the IQ distribution, one in four children may be misclassified if only a short form is used. Clearly, short forms should not be used when important clinical decisions concerning individual children have to be made.

Note that we are *not* recommending that the particular four or five subtest short forms examined in this paper should be used routinely for either research or screening for clinical purposes. These emerged in these particular samples as the subsets which accounted for most variance in FSIQ. As Kaufman (1979) cogently argues, in selecting a short form the purposes of the investigation have to be borne in mind. Usually, clinicians will want to tap the Verbal and Performance scales equally.

Indeed, the difficulty of relying solely on statistical selection is well illustrated in Table 1. In the London sample, Object Assembly emerges as a good predictor, but it takes the longest time of all subtests to administer. Where considerable time saving is paramount, this short form is not to be recommended. The contrast among the sets of tests selected in the London/Leeds and the Beck *et al.* studies underlines the point made by Silverstein (1975), namely that many combinations of four or five subtests yield good predictions of FSIQ on the WISC-R.

Short forms of WISC-R will continue to be used. This study demonstrates that such abbreviated tests are more robust than many critics assume in so far as there is good evidence for little loss of power in cross-validation. However, care must be taken to ensure that appropriate equations are selected to suit the population from which the subject is drawn and the possibility of misclassifying scores at the extremes of the range must be borne in mind. Short forms can be recommended for most research and screening purposes, but fuller testing is recommended for individual clinical purposes.

ACKNOWLEDGEMENTS — These studies were supported by grants from the Medical Research Council and the Commission of the European Communities to WY and RL. We gratefully acknowledge the assistance of officials and teachers of the Inner London, Bexley and Leeds Education Authorities.

Correspondence and requests for reprints should be addressed to Professor W. Yule, Institute of Psychiatry, de Crespigny Park, London, SE5 8AF.

REFERENCES

- BECK, N. C., RAY, J. S., SEIDENBERG, M., YOUNG, T. R., and GAMACHE, M. P. (1983). Development and cross-validation of a short form for the WISC-R. *J. consult. clin. Psychol.*, 51, 864-869.
- HUNTER, J., URBANOWICZ, M. A., YULE, W., and LANSDOWN, R. (1985). Automated testing of reaction time and its association with lead in children. *Int. Arch. occup. environ. Health*, 57, 27-34.
- KAUFMAN, A. S. (1977) Comments: Should short form validating coefficients be corrected? *J. consult. clin. Psychol.*, 45, 1159-1161.
- KAUFMAN, A. S. (1979) *Intelligent Testing with the WISC-R*. New York: Wiley.
- LANSDOWN, R., YULE, W., URBANOWICZ, M. A., and HUNTER, J. (1986). The relationship between blood lead concentration, intelligence, attainment and behaviour in a school population: the second London study. *Int. Arch. occup. environ. Health*, 57, 223-235.
- PHILLIPS, C. J. (1984). Use of abbreviated cognitive tests in the assessment of children's special educational needs. *Br. J. educ. Psychol.*, 54, 163-176.
- QUATROCCHI, H., and SHERRETS, S. (1980). WISC-R: the first five years. *Psychology in the Schools*, 17, 297-312.
- SILVERSTEIN, A. B. (1975). Validity of WISC-R short forms. *J. clin. Psychol.*, 31, 696-697.
- SILVERSTEIN, A. B. (1982). Validity of Satz-Mogel-Yudin-Type short forms. *J. consult. clin. Psychol.*, 50, 20-21.

YULE, W., LANSDOWN, R., MILLAR, I. B., and URBANOWICZ, M. A. (1981). The relationship between blood lead concentrations, intelligence and attainment in a school population: a pilot study. *Dev. Med. Child Neurol.*, 23, 567-576.

(Manuscript received 18th August, 1988)

Br. J. educ. Psychol., 59, 372-377, 1989

A NUTRITION THEORY OF THE SECULAR INCREASES IN INTELLIGENCE; POSITIVE CORRELATIONS BETWEEN HEIGHT, HEAD SIZE AND IQ

BY RICHARD LYNN

(University of Ulster, Coleraine, Northern Ireland)

SUMMARY. The thesis is advanced that the secular increases in intelligence over the last half century are largely due to improvements in nutrition. These have brought about increases in height, head size and brain size of approximately one standard deviation, about the same magnitude of increase as has taken place in intelligence. The thesis requires the existence of a positive correlation between head size and intelligence. Data for 310 9-10 year-olds obtained a correlation of +0.21.

INTRODUCTION

Recent studies in a number of economically developed nations including Britain, the United States, Australia, New Zealand and several countries of Continental Europe have shown that the intelligence of children has increased by approximately three IQ points per decade or one standard deviation over the last half century (Lynn and Hampson, 1986, Flynn, 1987). The magnitude of the increase has surprised students working in this field. One leading investigator has concluded that the increase cannot be genuine and hence that intelligence tests do not measure intelligence (Flynn, 1987). This conclusion is not accepted. The construct and predictive validity of intelligence test, recently reviewed by Gottfredson (1986), is too securely established to reject the tests as measures of intelligence. In this paper it is proposed that the increase is genuine and that the major factor responsible has been improvements in nutrition. The objective of the paper is to present this theory and report some data consistent with it concerning positive associations between intelligence, height and head size.

There are two principal arguments for the nutrition theory of the secular increases in intelligence. Firstly, there is direct evidence that nutrition affects intelligence; and secondly, nutrition has improved in the economically developed nations over the course of the last half century. On the basis of these two propositions it is not unreasonable to infer that the improvements in nutrition have played a part in the increases in intelligence. The evidence for the two propositions is now summarised.

The thesis that nutrition affects intelligence obtains support from various kinds of evidence. A number of studies have demonstrated that malnourished children tend to have low IQs but in many of these it is difficult to control for confounding factors such as low parental intelligence, poverty and other adverse conditions associated with malnourishment.

One of the more persuasive studies is that of Winick *et al.* (1975) of Korean infants adopted by American parents. One hundred and eleven Korean female babies were classified into three groups of malnourished, moderately nourished and well nourished on the basis of their height and weight. They were placed with American adoptive parents before the age of three years. The mean IQs of the three groups at the age of around 10 years were 102 (malnourished), 106 (moderately nourished) and 112 (well nourished), the difference between groups 1 and 3 being statistically significant. It is difficult to see how the results can be plausibly explained except in terms of a permanently adverse effect of poor nutrition in infancy on subsequent intelligence.

Probably the most convincing evidence for an effect of nutrition on intelligence comes from studies of identical twins with different birth weights. Where this occurs, the different birth weights are due to insufficient nutrition reaching one foetus from the mother's placenta, causing retarded growth and light birth weight. The first study to investigate a possible effect of low birth weight on later intelligence was carried out by Churchill (1965) and

Willerman and Churchill (1967). They reported data for 27 monozygotic twin pairs with differing birth weights intelligence-tested with the WISC at a mean age of 9.6 years. The performance IQ of the lighter twins was 5.3 points lower than that of the heavier (statistically significant) and the verbal IQ 0.4 points lower. A recent study from Denmark reports similar results for 14 monozygotic twins given the WISC at a mean age of 13 years. The performance IQ of the lighter twin was 7.1 points lower than that of the heavier although there was no difference on the verbal IQ (Hendrichsen *et al.*, 1986). The use of monozygotic twins in these studies controls for genetic differences and postbirth environmental differences and so the results are advanced as convincing evidence that sub-optimal nutrition at an early stage has a permanent adverse effect on subsequent intelligence.

The second premise of the nutrition theory of the secular increase in intelligence is that nutrition has improved over the course of the last half century. In the period between the two world wars several nutrition surveys showed that sub-optimal nutrition was prevalent for substantial proportions of the population. In Britain it was estimated that less than half the population were receiving adequate intakes of vitamins and minerals (Corry Mann, 1926; Orr, 1936). Similar conclusions were reached in the United States and Japan (Palmer, 1935; Takahashi, 1966).

In the post-World War 2 decades, rising living standards have enabled people to buy more nutritious foods. The result has been that the height of children and young adults in the economically developed nations has increased over the last half century by around 7.8 cm. This represents a rise of approximately one standard deviation, the same magnitude of increase as has taken place in intelligence (Van Wieringen, 1978; Whitehead and Paul, 1988).

The increases that have taken place in height have also occurred in head size. These increases have been recorded among infants, children and young adults. In Britain the head circumference of infants and young children has increased by approximately 1.5-2.0 cm over the last half century (Whitehead and Paul, 1988). Similar results have been obtained in Hong Kong (Davies *et al.*, 1985). These increases in head size are also of the order of one standard deviation over a half century. Head size is correlated at a magnitude of approximately 0.8 with brain size (Brandt, 1978). Thus the secular increases in head size represent increases in brain size.

The nutrition theory of the secular increases in intelligence proposes that the increases in head size are causally associated with the increases in intelligence. For this to be the case, there would have to be a positive correlation between head size and intelligence. The present paper reports data on this correlation and is therefore a test of the theory.

A number of recent writers have rejected the proposition that there is any association between head size and intelligence. Thus "it is generally accepted that in healthy subjects there is little or no correlation between brain size and mental performance" (Engsner, 1974, p. 37); and "despite the terms egghead, pinhead and bonehead, there is really no evidence to show that brain size is positively correlated with higher intelligence" (Latham, 1974, p. 549). There were, however, several studies reporting the existence of such a correlation in the early years of the century. Pearson (1906) obtained a correlation of +0.11 between head circumference and the ability of children as rated by teachers; Pearl (1906) obtained a correlation of +0.14 between head circumference and ratings of ability on Bavarian soldiers; and Murdoch and Sullivan (1923) reported a correlation of +0.22 between head diameter and IQs among 600 American children. Nevertheless, there is a tendency for the credibility of antiquated findings such as these to deteriorate with age and many contemporary scholars are either unaware of the early studies or dismiss them.

Before describing the study, it should be made clear that there are two mechanisms through which the secular increases in head size could be causally related to the increases in intelligence. The first is that larger heads contain larger brains, and it may be that larger brains tend to be more intelligent. This hypothesis would be consistent with the trend across species for large brained animals to be more intelligent than small brained animals (Jerison, 1982; Mackintosh *et al.*, 1985). It is not unreasonable that the same relationship may hold within humans.

There is, however, a second hypothesis that improvements in nutrition may have exerted their effects on intelligence through effects on the neurological development of the brain, probably by generating improvements in the growth of the number of glial cells, the myelination of the neurons, the growth of dendrites and the formation of synaptic connections. There is evidence from animal studies and from autopsies on humans that malnutrition adversely affects these neurological developments (Winick *et al.*, 1970; Dobbing, 1984). The effect of nutrition on intelligence could operate solely through the internal neurological development of the brain and its effect on increasing brain size could be a correlate of increased intelligence but not a cause.

Which of these two mechanisms is operative is not crucial for the nutrition theory of the secular increases in intelligence. The theory only predicts that there should be positive associations between height, head size and intelligence, and it is to the test of this prediction that we now turn.

METHOD

The 310 subjects were 9- and 10-year-old children at four primary schools in the small town of Coleraine in Northern Ireland (161 boys and 149 girls). All the children in the grades for this age range were tested, except for absentees on the days of the testing. All the children were Caucasian.

The intelligence test used was the Primary Mental Abilities (Thurstone, 1963). This test provides IQs for the reasoning, verbal, numerical, spatial and perceptual speed primary abilities, and in addition a total IQ. Head circumference was measured at its maximum, and height and weight were also recorded.

RESULTS

Descriptive statistics giving means and standard deviations for boys and girls separately are presented in Table 1. It will be seen that the girls obtained significantly higher means for the number, reasoning and perceptual speed primaries. These are typical results for British 9-10 year-olds. It will also be seen that the means are higher than the American norms. The mean total IQ for boys and girls together is 109.82. The principal reason for this high mean is that the American norms were obtained about 1960 (the test was published in 1963) whereas the present Northern Ireland data were obtained in 1987. Mean IQs in the United States have been increasing by approximately three IQ points per decade, as noted in the introduction to this paper. Hence the contemporary American mean on the PMA can be assumed to be about 108 and very close to the mean for our Northern Ireland sample. The means for height, weight and head circumference are normal.

TABLE 1
MEANS AND STANDARD DEVIATIONS FOR IQs, HEIGHT, WEIGHT AND HEAD CIRCUMFERENCE

	MALES		FEMALES		t	P
	Mean	SD	Mean	SD		
V IQ	109.70	18.39	110.30	16.47	0.30	NS
N IQ	105.50	14.41	108.93	13.31	2.17	0.031
S IQ	105.07	17.43	102.36	14.47	1.48	NS
R IQ	105.16	16.47	109.99	15.64	2.64	0.009
PS IQ	105.10	16.71	110.05	15.90	2.67	0.008
Tot IQ	108.63	16.50	111.01	18.05	1.21	NS
Age-Months	125.35	7.20	125.36	6.97	0.01	NS
Height-cm	141.34	6.90	140.34	6.74	1.29	NS
Weight-kg	33.42	6.32	34.67	7.90	1.55	NS
Head c-cm	55.07	1.77	54.84	1.68	1.20	NS

The product moment correlations between head circumference and the IQs for the entire sample were 0.14** (reasoning), 0.26*** (verbal), 0.11* (numerical), 0.12* (spatial), 0.07 (perceptual speed) and 0.18*** (total IQ). The asterisks denote statistical significance at the 5 per cent, 1 per cent and 0.1 per cent levels.

The complete correlation matrix for boys and girls separately is given in Table 2. It will be seen that the five primary abilities are quite highly intercorrelated, as would be expected. The IQs are positively correlated with height: all six correlations are statistically significant for girls and three are statistically significant for boys. Weight shows no correlation with IQs or possibly a very small negative correlation among girls. Head circumference is positively correlated with IQ for each sex separately, five of the six correlations being statistically significant among girls and two of the six among boys.

TABLE 2
CORRELATION MATRIX FOR BOYS (BOTTOM LEFT) AND GIRLS (TOP RIGHT)

	Height	Weight	Head C	V	N	S	R	PS	Total IQ
Height	—	63	38	19	26	16	20	30	27
Weight	73	—	35	-12	-10	-09	-18	-03	-13
Head C	26	36	—	28	23	25	19	12	23
IQ—V	16	07	25	—	67	54	70	63	84
IQ—N	04	-03	03	62	—	48	70	69	81
IQ—S	11	11	02	44	43	—	55	48	68
IQ—R	12	-01	12	70	55	54	—	65	82
IQ—PS	19	10	05	59	70	45	58	—	83
IQ—									
Total	19	08	15	86	81	64	83	83	—

Decimal points omitted. Statistical significance levels are 13 ($P < 0.05$) and 17 ($P < 0.01$).

Since the association between brain size and cognitive ability across species normally takes body size into account by the derivation of an encephalisation quotient expressing brain size in relation to body size, it may be thought appropriate to consider this question in relation to the present sample. The simplest measure of body size is weight, but since this has no correlation with IQ, there is no effect to be removed by partialling it out. It may be arguable that height should be partialled out and for those who take this view the partial correlations between head size and IQ, partialling out height, are +0.07 (boys) and +0.14* (girls). However, it is considered that these are not appropriate calculations because height is not causal to head size. Rather, the positive correlation between height and head size arises because of a common effect of nutrition.

All of the 12 correlations between head circumference and IQ for the primary abilities for boys and girls are positive. Six of the 12 correlations calculated on boys and girls separately are statistically significant and on the combined sample five out of the six correlations are statistically significant. The overall correlation on both sexes between total IQ on the test and head circumference of +0.18 ($P < 0.01$) can appropriately be corrected for unreliability of the intelligence test and of head size as a measure of brain size. Assuming reliability coefficients of 0.9 and 0.8 respectively, the corrected correlation is +0.21.

DISCUSSION

It is considered that the findings of significant positive associations between height, head size and intelligence among this child population provide corroboration for the nutrition theory of the secular increases in intelligence. It is not easy to explain this set of intercorrelations except in terms of the theory that nutrition affects all three variables and hence brings them into positive correlation. This in turn implies that nutrition is a significant determinant of intelligence among contemporary children.

The nutrition theory of the secular rise of intelligence provides an explanation for one of the puzzling features of these increases. This is that the visuo-spatial abilities have shown greater rises than the verbal-educational abilities. When the Wechsler tests have been used, greater increases have been found on the performance scale, largely measures of visuo-spatial abilities, than on the verbal scale. These differential rates of increase have been found in the United States, Japan, Austria, France and West Germany (Flynn, 1987). Similarly in Britain the increases on Cattell's Culture Fair test and on the Coloured Progressive Matrices, non-verbal tests with a visuo-spatial component, have been approximately 2.5 IQ points per decade, as compared with the smaller increases of approximately 0.4 IQ points per decade on the Mill Hill Vocabulary Test and 1.1 IQ points per decade on the verbal-educational test used in the 1932 Scottish survey (Lynn and Hampson, 1986; Lynn *et al.*, 1987, 1988). These differential rates of increase of the two abilities are probably surprising, since it might be expected that with improvements in education and affluence the verbal-educational abilities would have shown greater rises than the visuo-spatial. The nutrition theory of the increases explains this problem. It was noted in the introduction that the studies of the intelligence of identical twins with different birth weights show that the visuo-spatial abilities are more seriously impaired in light birth weight babies while the verbal abilities are hardly affected. This indicates that, for some reason not at present understood, sub-optimal nutrition has more serious adverse effects on the visuo-spatial abilities. This would explain why, as nutrition has improved over the course of the last half century, it is the visuo-spatial abilities which have shown the greater improvement.

If the nutrition theory of the secular increase in intelligence is correct it carries important implications for educational and developmental psychology. During the last 20 years or so considerable efforts have been expended to attempt to raise the intelligence of deprived children through headstart programmes. These have largely concentrated on providing cognitive stimulation for young children in accordance with the widespread belief among educational and developmental psychologists that cognitive stimulation is the principal environmental determinant of intelligence. Yet the results of the headstart programmes have been disappointing and although they may have increased specific cognitive skills for two or three years it is doubtful whether they have had any permanent effect on intelligence (Bronfenbrenner, 1974; Jensen 1989). Yet in spite of the disappointing results of these attempts to raise children's intelligence by increased cognitive stimulation, intelligence has been increasing spontaneously over recent decades. Nutrition theory suggests that this may be because it is nutrition rather than cognitive stimulation which is the major environmental determinant of intelligence. This in turn suggests that the way to tackle the problem of low intelligence among deprived children may be through measures designed to improve their nutrition.

ACKNOWLEDGMENTS. I am indebted to Miss Susan Clark for assistance in the collection of data.

Correspondence and requests for reprints should be addressed to Professor Richard Lynn, University of Ulster, Coleraine, Co. Londonderry, BT52 1SA, Northern Ireland.

REFERENCES

- BRANDT, I. (1978). Growth dynamics of low-birth weight infants with emphasis on the perinatal period. In FALKNER, F., and TANNER, J. M. (Eds.), *Human Growth*, Vol. 2. New York: Plenum Press.
- BRONFENBRENNER, U. (1974). Is early intervention effective? *Teachers Coll. Rec.*, 76, 279-303.
- CHURCHILL, J. A. (1965). The relationship between intelligence and birthweight in twins. *Neurol.*, 15, 341-347.
- CORRY MANN, H. C. (1926). *Special Reports Series 104-105*. London: Medical Research Council.
- DAVIES, D. P., LEUNG, S. F., and LAU, S. P. (1985). Secular trends in head growth. *Arch. Dis. Child.*, 60, 623-624.
- DOBING, J. (1984). Infant nutrition and later achievement. *Nutr. Rev.*, 42, 1-7.
- ENGSTNER, G. (1974). Brain growth and motor nerve conduction velocity in children with protein-calorie malnutrition. *Acta Univers. Upsal.*, 1-60.
- FLYNN, J. R. (1987). Massive IQ gains in 14 nations: what IQ tests really measure. *Psychol. Bull.*, 101, 171-191.

- GOTTFREDSON, L. S. (1986). Societal consequences of the 'g' factor in employment. *J. voc. Behav.*, 29, 379-410.
- HENDRICHSEN, L., SKINHOI, K., and ANDERSEN, C. E. (1986). Delayed growth and reduced intelligence in 9-17 year old intrauterine growth retarded children compared with their monozygous co-twins. *Acta Paediat. Scand.*, 75, 31-35.
- JENSEN, A. R. (1989). Raising IQ without increasing g? *Dev. Rev.*, (to appear).
- JERISON, H. J. (1982). The evolution of biological intelligence. In STERNBERG, R. J. (Ed.), *Handbook of Human Intelligence*. Cambridge: Cambridge University Press.
- LATHAM, M. C. (1974). Protein-calorie malnutrition in children and its relation to psychological development and behaviour. *Physiol. Rev.*, 54, 541-565.
- LYNN, R., and HAMPSON, S. (1986). The rise of national intelligence: evidence from Britain, Japan and the United States. *Person. indiv. Diff.*, 7, 23-32.
- LYNN, R., HAMPSON, S., and MULLINEUX, J. C. A long term increase in the fluid intelligence of English children. *Nature*, 328, 797.
- LYNN, R., HAMPSON, S., and HOWDEN, V. (1988). The intelligence of Scottish children 1932-1986. *Stud. Educ.*, 6, 19-25.
- MACKINTOSH, N. J., WILSON, B., and BOAKES, R. A. (1985). Differences in mechanisms of intelligence among vertebrates. *Phil. Trans. Roy. Soc. London*, 8, 308, 53-65.
- MURDOCH, J., and SULLIVAN, L. R. (1923). A contribution to the study of mental and physical measurements in normal school children. *Am. phys. Educ. Rev.*, 28, 209-215, 276-280, 328-330.
- ORR, J. B. (1936). *Food, Health and Income*. London: Macmillan.
- PALMER, C. E. (1935). Height and weight of children of the depressed poor. *Public Health Reports*, 59, 33-35.
- PEARL, R. (1906). On the correlation between intelligence and the size of the head. *J. comp. Neurol. Psychol.*, 16, 189-199.
- PEARSON, K. (1906). On the relationship of intelligence to size and shape of head, and to other physical and mental characters. *Biometrika*, 5, 105-146.
- TAKAHASHI, E. (1966). Growth and environmental factors in Japan. *Hum. Biol.*, 38, 112-130.
- THURSTONE, T. G. (1963). *Primary Mental Abilities*. Chicago: Science Research Associates.
- VAN WIERINGEN, J. C. (1978). Secular growth change. In FALKNER, R., and TANNER, J. M., (Eds.), *Human Growth*, Vol. 2. New York and London: Plenum Press.
- WHITEHEAD, R. G. and PAUL, A. A. (1988). Comparative infant nutrition in man and other animals. In WHITEHEAD, R. G., and PAUL, A. A. (Eds.), *Proceedings of the International Symposium on Comparative Nutrition*. London: J. Libbey.
- WILLERMAN, L., and CHURCHILL, J. A. (1967). Intelligence and birth weight in identical twins. *Child Dev.*, 38, 623-629.
- WINICK, M., ROSS, P., and WATERLOW, J. (1970). Cellular growth of the cerebrum, cerebellum and brain stem in normal and marasmic children. *J. exp. Neurol.*, 26, 393-400.
- WINICK, M., MEYER, K. K., and HARRIS, R. C. (1975). Malnutrition and environmental enrichment by early adoption. *Science*, 190, 1173-1175.

(Manuscript received 23rd November, 1988)

Br. J. educ. Psychol., 59, 379-388, 1989

CROSS-CULTURAL TRANSFER OF NON-VERBAL INTELLIGENCE TESTS: AN (IN)VALIDATION STUDY

By RENE S. PARMAR

(*State University of New York at Buffalo, USA*)

SUMMARY. Non-verbal intelligence tests have been used widely in India for measurement and research without adequate examination of their validity in the Indian culture. The present study provides evidence that a test meeting the specified conditions for "culture-fairness" does not demonstrate the expected equal performance across cultures. A comparison of children from the normative sample of the *Test of Non-Verbal Intelligence* and a matched sample of Indian children was conducted. Confirmatory factor analysis was not supported by the data structure. Factor scales comparisons between the two groups indicated significant differences. Item analyses indicated that items requiring matching were those significantly different. Comparison of students at different levels of ability and aptitude within the Indian sample indicated that the test did not discriminate sufficiently for classification or diagnostic usefulness.

INTRODUCTION

In recent years there has been a growing recognition in India of the need to develop psycho-educational testing measures which demonstrate validity in the Indian culture. Though research exists that is conventional and sound, most of it is directed towards American and Western European needs, rather than problems and trends in India (Sinha, 1973). Thus, there are insufficient data available that could guide policy-making for educational practice in India.

An emerging educational interest in India, special education, has emphasised the need for reliable and efficient instruments to assess cognitive capacity. Major tests currently used in India for intellectual assessment include (a) Raven's Progressive Matrices (RPM) (Raven, 1960), (b) Wechsler Intelligence Scale for Children (WISC) (Wechsler, 1949), (c) Non-Language Test of Verbal Intelligence (Chatterji and Mukherjee, 1967), (d) Bhatia's Battery (Bhatia, 1955), (e) Kerala Non-Verbal Test of Intelligence (Nair, 1968), (f) Cattell Culture-Free Intelligence Test (CCFT) (Cattell and Cattell, 1959), and (g) Goodenough Draw-A-Man Test (Goodenough, 1926). These are adaptations primarily from western psychological thought, and frequently bear little relevance to conditions in India. Further, they do not address the vast differences in socio-economic status, literacy, and language that exist among the various subcultures in India. Most need extensive adaptations, to the extent that the adapted test does not completely reflect the theoretical basis on which it was formed. For example, the Information subtest of the WISC is simply deleted when testing Indian subjects, and this scale is not considered when computing IQ scores. Further, verbal and non-verbal or performance tests are used interchangeably, which is likely to bias predictions based on test results (Nair, 1975). This issue has serious negative implications for educational or clinical decisions based on results.

Efforts to use standardised tests in comparative studies across cultures have resulted in confusing and often contradictory results and frustration. One explanation for this lack of acceptance of "culture-fair" measures may be that, in studies of intelligence testing and consequent conclusions about cognitive ability, it has rarely been acknowledged that the instruments used in such studies were thoroughly inadequate for the culture under study (Zaidi, 1979). Further, limited attempts at cross-cultural study have not always been guided by sound theoretical orientation taken from main psychological literature.

Table 1 presents a list of validity studies conducted with non-verbal intelligence measures. Some of the problems of current research programmes are:

- (1) A majority of studies have used only one method to investigate the applicability of the test under consideration.

TABLE 1

VALIDITY STUDIES ON NON-VERBAL TESTS IN INDIA — METHODOLOGY

Raven's Progressive Matrices

Author	Method	Comparison	Test	Findings
Desai (1980)	F. A.	Vbl. & N-Vbl.	RPM	not supportive
Dosajh (1958)	Corr.	N-Vbl. & Ach.	RPM	supportive
Ghuman (1975)	ANOVA	Br. & Ind.	RPM	not supportive
Mehrotra (1967)	t	Vbl. & N-Vbl.	RPM	not supportive
Mohan (1972)	Corr.	M & F	RPM	supportive
Mohan (1972)	Corr.	Vbl. & N-Vbl.	RPM	supportive
Mohanty (1980)	ANOVA	High & Low SES	RPM	not supportive
Mohsin (1959)	Corr.	Ach. & N-Vbl.	RPM	not supportive
Patted (1967)	Corr.	Vbl. & N-Vbl.	RPM	not supportive
Rath (1954)	ANOVA	Vbl. & N-Vbl.	RPM	supportive
M. Sinha (1975)	X & SD	Age levels	RPM	supportive
Sinha & Chandrakala (1972)	Corr.	N-Vbl. & N-Vbl.	RPM	not supportive

Cattell Culture-Fair Test of Intelligence

Author	Method	Comparison	Test	Findings
Rao (1965)	X & SD	Vbl. & N-Vbl.	CCFT	not supportive
Singh & Hundal (1971)	X & SD	Vbl. & N-Vbl.	CCFT	supportive

Draw-A-Man

Author	Method	Comparison	Test	Findings
Bevli (1982)	ANOVA	High & Low SES.	DAM	not supportive
S. N. Rao (1962)	Corr.	N-Vbl. & Ach.	DAM	not supportive

Other

Author	Method	Comparison	Test	Findings
Chatterji & Mukherjee (1967)	X & SD	Vbl. & N-Vbl.	CMNV	supportive
Chawla (1969)	Corr.	Reliability	Kahn	supportive
Deb (1966)	Corr.	Reliability	GIT	supportive
Hundal (1965)	F. A.	Vbl. & N-Vbl.	BB	supportive
Kakkar (1975)	t	Lang. & N-Vbl.	LT	not supportive
A. S. Nair (1973)	F. A.	N-Vbl. Tests	KNV	supportive
K. S. Nair (1975)	F. A.	Vbl. & N-Vbl.	KNV	supportive
Rao & Gupta (1984)	X & SD	Validity	AH5	supportive
R. R. P. Sinha (1980)	Corr.			
Warhadpande & Sethi (1964)	t	Tribal & Non-Tribal	APS	supportive
	t	Lang. & Ach.	R-T	supportive

Note: F. A. = Factor Analysis; Corr. = Correlation; X & SD = Means and Standard Deviations; Vbl. = Verbal tests; N-Vbl. = Non-verbal tests; M = Males; F = Females; Br. = British; Ind. = Indian; SES = Socio-economic status; Lang. = Language tests; Ach. = Achievement tests; CMNV = Chatterjee & Mukherji Non-Verbal Test; GIT = General Intelligence Test; LT = Leiter test; KNV = Kerala Non-Verbal Test; APS = Alexander Performance Scale; R-T = R Test.

- (2) Several studies utilise purely descriptive methodology, yet make inferential statements when describing results.
- (3) A majority of studies look at only one dimension of reliability or validity.
- (4) The studies provide conflicting evidence regarding test validity.
- (5) Several studies have employed obscure tests with unknown norms.

The purpose of this study is to investigate the extent to which a non-verbal test of intelligence, the *Test of Non-Verbal Intelligence* (TONI) (Brown *et al.*, 1982) may be used for assessing intellectual abilities of children in India. This investigation is important to both Indian, as well as other third world, educators and psychologists: current instruments in wide use throughout Asia and Africa do not demonstrate adequate validity, as procedures for development and cultural transport are frequently not in adherence to recommended guidelines for such practice.

Two common assumptions in cross-cultural test transferability are that (a) experts are adequate judges of whether or not a test is likely to be biased, and (b) it is possible, based on intuitive logic, to determine which items will reveal bias. These assumptions have not always been empirically validated. The present study provides evidence that popular assumptions do not hold up against close scrutiny.

The TONI was selected for use in the present study for several reasons. First, it meets criteria specified by Jensen (1980) for selection of culturally-reduced tests, i.e., (a) it is a performance measure, (b) instructions are presented in pantomime to subjects, (c) it provides preliminary practice items, (d) it is untimed, (e) items are comprised of abstract figural content, (f) items require abstract reasoning rather than factual information, and (g) problems are designed in such a way that subjects may not guess at answers from memory of similar items encountered in the past. Due to the nature of the instrument, it can be used in both cultures in its original standardised form which means that the results are accessible to a variety of analyses on equivalence (Malpass and Poortinga, 1986).

Second, the TONI is a relatively recently developed test. Tests commonly used for cross-cultural study have been developed and standardised several decades ago. Norms developed at that time would not be considered representative today (D. Sinha, 1981). Further, the standardisation sample itself, in many cases, has not been representative even of the culture in which the test was originally developed. On the other hand, the standardisation sample of the TONI appears to have been carefully chosen to be representative of the relevant American population (Clark, 1985).

The authors of the TONI developed the test to fill the need for an instrument which measures intellectual ability of persons for whom traditional tests, which use written or spoken language as part of their content or testing format, are inappropriate. These include "people who are unable to read or write and people who have poor or impaired language skills, such as aphasics, non-English speakers, and individuals who are mentally retarded, learning disabled, deaf, or culturally different" (Brown *et al.*, 1982, p. 2).

METHOD

Design

To control for extraneous variance and eliminate rival hypotheses, cross-cultural studies must employ controls and procedures that are not generally applied to experimental research. This is because cross-cultural research seldom meets the requirement of the experimental paradigm, such as (a) equating samples through randomisation (Campbell and Stanley, 1966, p. 2), and (b) exercising control over all treatments under study (Cook and Campbell, 1979, p. 98-99) and, therefore, cannot use the inferential strategy of a proper experiment (Malpass and Poortinga, 1986). Consequently, statistical techniques must be utilised to control for extraneous effects.

Prior research in cross-cultural transferability of tests has provided evidence that observing certain methodological constraints leads to results that may be interpreted with greater confidence than those obtained from traditional experimental study. The major areas of differentiation between cross-cultural and traditional experimental research are

(a) sample selection, (b) instrument selection, (c) data collection procedures, and (d) statistics for equivalence.

Problems in analysing cross-cultural equivalence of tests arise when only one strategy or technique is used. A multistrategy approach is recommended (Hui and Triandis, 1983, 1985) which combines highly mathematical and statistical techniques with those that place considerable demands on the researcher's conceptualisation ability.

Sample selection

Based on a review of studies previously conducted in India (cf. Nanda, *et al.*, 1965; Jachuck and Mohanti, 1974; Panda and Das, 1970; Das and Singha, 1975; D. Sinha, 1978), the following variables were selected to be controlled: (a) literacy level, (b) area of residence, and (c) socio-economic level.

(a) *Literacy level.* Students selected for the present study were 7-, 8- and 9-year-old school children currently enrolled in second ($N = 36$) and third ($N = 54$) grade classes. Ninety normally-achieving children were selected from schools in Lucknow, India. This number was considered adequate for the statistical procedures to be applied to the data. Eighteen mentally retarded children selected for the study were also currently enrolled in school. Since no national criteria exist to identify the mentally retarded, children selected were those enrolled in a school for the mentally retarded, and whose WISC-R scores indicated their cognitive aptitude to be between 69 and 40 IQ points. The comparison group selected from the TONI normative sample consisted of children matched on grade level with no significant difference in mean age from the Indian normally-achieving group ($t = 1.08$, $df = 224$, $P < 0.01$). Table 2 presents mean ages of students and the age range.

TABLE 2
MEAN AGE AND AGE RANGES OF SUBJECTS IN MONTHS

Group	Mean Age	Range
Indian (Normally-Achieving)	95.9	84-120
Indian (Mentally Retarded)	108.1	84-130
American (Normative Sample)	95.0	84-111

(b) *Area of residence.* Children for the present study were selected from schools in a major urban area (Lucknow, India) where they have exposure to a variety of experiences. Children selected from the normative sample of the TONI were also from urban residential areas.

(c) *Socio-economic level.* Socio-economic variables were controlled for by selection of children from middle to upper-middle class backgrounds, enrolled in private schools. In India private school education is accessible only to middle- and upper-middle-class children, and is comparable to standards of education in the US. In addition, schools where the medium of instruction is English were selected, where children are exposed to western literature, art, and concepts. Hollingshead's Index was used to verify socio-economic standing. Table 3 presents information on socio-economic background of the selected groups.

Instrumentation

The TONI is described by its authors as a language free test designed to measure cognitive ability, appropriate for use with subjects ranging in age from 5:0 through 85:11 years. The basis of all the TONI items is visual problem-solving (Brown *et al.*, 1982). The authors

TABLE 3
PERCENTAGES OF SUBJECTS AT HOLLINGSHEAD INDEX SOCIO-ECONOMIC LEVELS

Level	Indian (Normally- Achieving)	American (Normative- Sample)
I	28.1	16.0
II	40.4	23.2
III	31.5	48.8
IV	—	12.0

selected this test in the light of the work of Sternberg (1980) who considered problem-solving to be a general component of intelligent behaviour rather than a subskill. Further, Resnick and Glaser (1976) have also cited problem-solving as the basis for functional independence. Problem-solving lends itself readily to the abstract content and the non-verbal testing format used in the TONI.

The TONI purports to measure problem-solving ability, rather than general "intelligence", by measuring respondents' performance on five types of problem-solving tasks:

1. Simple matching.
2. Analogies: (a) matching, (b) addition, (c) subtraction, (d) alteration, and (e) progressions.
3. Classification.
4. Intersections.
5. Progressions.

The TONI was standardised on 1,929 persons ranging in age from 5 years 0 months to 85 years 11 months. Demographic characteristics approximated the 1980 United States population (47 per cent male, 53 per cent female; 77 per cent Caucasian, 14 per cent Negroid, 9 per cent Other; 78 per cent urban, 22 per cent rural).

Internal consistency analysis revealed coefficient alphas of 0.78 to 0.91. Kuder-Richardson 21 coefficients of 0.8 and 0.9 are reported for two age groups. K-R 21 coefficients computed on mentally retarded, deaf, and learning disabled samples consistently exceeded 0.8.

Concurrent and construct validity was established through comparison with Raven's Progressive Matrices, Leiter International Performance Scale, WISC-R, Otis-Lennon Mental Ability Test, Iowa Test of Basic Skills, SRA Achievement Series, and the Stanford Achievement Test. All coefficients reported exceed 0.35, with 41 per cent exceeding values of 0.80. Further, mean raw scores of the sample follow developmental patterns reported in current psychological literature.

Data collection procedures

Procedures which control for extraneous effects as described by Cronbach and Drenth (1972), Irvine and Carroll (1980), and D. Sinha (1981) were utilised. Briefly they are:

- (1) Examiner effect. To maximise correct interpretation of response, examiners trained in education and psychology, indigenous to the children's culture were used.
- (2) Setting effect. Students were tested at their own schools, in as close proximity to their classrooms as possible.
- (3) Effect of unfamiliar materials. The TONI provides practice items, so testees may familiarise themselves with the format of the test and the type of response desired.
- (4) Response time constraints. The TONI is untimed, allowing sufficient time to make a response without undue pressure for speedy performance.

(5) Ambiguous directions. Since pantomime is used rather than written or oral language, the risk of non-comprehension due to linguistic complexity are minimised. Administration of training items permits verification of comprehension of pantomimed directions.

Statistics for equivalence

Concern has been expressed about equivalence of data sets used in cross-cultural studies. The major types of equivalence described as essential by Hui and Triandis (1985) are addressed below.

(1) Conceptual/functional equivalence. The selection of students with similar levels of literacy in both cultures, and comparable exposure to testing contributed to maintaining functional equivalence. Since the TONI items consist of stimuli that are "not symbolic and have no inherent meaning" (Brown *et al.*, 1982), conceptual equivalence was maintained.

(2) Equivalence in construct operationalisation. As evidenced in the literature, the concept of intelligence is similarly operationalised in both India and the US. Indian psychology is rooted in American and European psychological thought, and in both problem-solving is considered a primary component of intelligence.

(3) Item equivalence. Utilisation of identical testing instruments in both cultures (India and the US normative group), made possible by the format of the TONI, ensured item equivalence.

(4) Metric/scalar equivalence. Since administration procedures and items were identical for both cultural groups, it was possible to use the same scoring methods and scales for direct comparisons.

RESULTS

Based on the definition of a scientific model presented by Reynolds and Brown (1984), a hierarchical procedure was used to determine the validity of the TONI. Content validity issues were first addressed by selection of instrumentation that meets criteria defined by Jensen (1980), Salvia and Ysseldyke (1981), Berk (1982), and Reynolds (1982) for a "culturally reduced" test. However, since such attempts in the past have failed to produce a test that was not culturally loaded in some way, further analysis was conducted.

Factor scale difference

In comparing factor scales derived through principal components analysis it was found that the American and Indian normally-achieving students evidenced significant differences on Scales 2, 3, 4, 5, and 6 (Table 4). No significant differences were found for Scale 1.

The difference in factor scales indicates differences in the underlying response pattern of the two groups on the hypothetical trait being measured (problem-solving). The scientific value of this difference is made more interpretable by the fact that many "ambient"

TABLE 4
FACTOR SCALE MEANS AND CRITICAL VALUES OF F

Group	Scale 1	Scale 2	Scale 3	Scale 4	Scale 5	Scale 6
Indian	0.01	2.88	1.21	0.02	0.52	0.21
American	0.15	5.10	1.41	0.09	0.82	0.36
F	2.37	98.42*	22.72*	11.95*	64.17*	41.53*
P	0.01	0.01	0.01	0.01	0.01	0.01

Note: An asterisk (*) indicates significant differences between the group means.

variables were accounted for. Ambient variables are defined as antecedents that are not considered within the context of the focal theory, but which can reasonably provide alternative explanations of observed differences (Poortinga and Malpass, 1986). In the present study, differences between the groups are not attributable to (a) variance in subject characteristics, and (b) variance in testing procedures as these ambient variables were controlled.

Differences in cognitive strategy application can be attributed to differences in educational training, cultural differences in strategy utilisation, or a different approach to the task of problem-solving. Further information on this dimension may be obtained by interviewing subjects and determining the reasons for their chosen responses. Such a qualitative approach would determine the strategies applied by subjects to the task. However, this procedure was beyond the scope of the present study: this would be an avenue for future research.

Thus the assumption may not be made that since the TONI is devoid of language and utilises non-representational symbols, it is "culture-fair". This supports the findings of Jensen (1968), Irvine (1969), Vernon (1969), Ortner (1972), Court (1982) and Gonzalez (1982) on other non-verbal measures of abstract problem-solving such as the CCFT, RPM, and the Leiter. It also supports the view proposed by Berk (1982), Reynolds (1982), and Reynolds and Brown (1984) that attempts to reduce cultural loading of aptitude test items have thus far resulted in failure.

Item characteristic curve differences

The item characteristic curves for the American and Indian normally achieving groups were compared. The difference between the average point biserial correlations for the two groups was 0.079. This difference was not significant at $P < 0.01$. For individual items, a comparison of point biserial correlation values evidenced that items 4, 8, 10, and 22 indicated significant differences. Deletion of these items would increase the discrimination power of the test.

Raw percentages of correct answers for each item were converted to a Rasch scale for comparison of the two groups. Item 4 was the only item indicating significant differences and it would thus be inappropriate for a test to be used with both Indian and American students. However, since this item represents only 0.05 per cent of the total number of items, it is doubtful whether deletion would have any significant effect on overall difficulty level of the test.

In examining items found to be biased in terms of their discrimination ability and difficulty level, it was found that differing items were predominantly those which required simple matching, or analogous matching strategies (Table 5).

TABLE 5
ITEMS EVIDENCING SIGNIFICANT DIFFERENCES BETWEEN INDIAN AND AMERICAN SUBJECTS

Item	Form A	Form B
1	simple matching	simple matching
3	simple matching	matching analogy
4	simple matching	matching analogy
5	matching analogy	classification
8	matching analogy	matching analogy
10	matching analogy	matching analogy
22	intersection	classification

This finding indicates that there appears to be some fundamental difference in the way young Indian and American literate children approach the task of matching to a given sample. It is not evident from the results whether this difference is due to the way the problem is presented, or whether the task of matching itself is approached differently, with the Indian group looking for a non-matching figure, rather than a matching one. Deletion of items requiring matching skills may provide a test that is more valid for use with Indian children.

It is not possible to infer that Indian children do not possess matching abilities, as evidence of their matching ability exists in that they have successfully acquired language and mathematical skills, both of which depend to some degree on the ability to match. The discovery of difference in a particular skill supports the findings of numerous other studies reviewed by Pick (1980). In the research discussed, differences in cognitive strategies of subjects are evident between two groups on a standardised test. However, both groups demonstrate the same level of ability when tested in a format more closely related to the demands of their cultural environment. Thus, claims of apparent cultural differences in ability to abstract or to think in generalities may not be made based on the results of a standardised test alone.

No indication of different approaches to tasks requiring matching have been reported in the literature on abstract problem-solving ability of Indian children. Further research with a larger sample would be necessary before conclusive statements may be made regarding matching skills of Indian children. No literature was found which examined item characteristic curves of Indian children on non-verbal tests which had been adapted from the West, so supportive statements could not be made.

Reynolds and Brown (1984) state that panels of expert judges in both minority and majority cultures have not, thus far, been able to predict any better than chance which items in any given aptitude test will prove to be biased against one of the cultures. The findings support their contention.

Achievement and ability level differences

To examine the discriminant validity, a comparison was made between the mean TONI Quotient scores of normally achieving ($X = 82.28$) and mentally retarded ($X = 60.30$) children in India. The 22-point difference between the two groups was highly significant ($F = 107.14$, $df = 1, 107$, $P = <0.01$). It represented a difference of more than one standard deviation (15 points), but less than two standard deviations, which is the conventional discrepancy for diagnosis of mental retardation in the US (AAMD, 1973).

Mean raw scores of third grade and second grade normally achieving Indian children were compared. The third graders ($X = 9.87$) scored significantly higher than the second graders ($X = 7.03$). An F value of 13.91 was obtained for 1 and 89 degrees of freedom, $P < 0.01$.

Further, when scores were converted to TONI Quotients, no significant differences were found, indicating that standard scores computed for the two groups reflect their different levels of achievement. Further study on children at different grade levels would be necessary to determine the discrimination ability of the TONI at various age and grade levels.

To determine predictive validity, multiple regression analysis of TONI Quotients of second and third graders, using their WISC scores as the criterion variable, was conducted. The intercept (α) was 110.25 for second graders, and 130.78 for third graders. Thus the TONI has different predictive meaning for the two groups in relation to the WISC. This indicates that the two groups differ on a third variable which correlates positively with both the test and the criterion. Grade level and age are two possible extraneous variables which may have had this effect.

Slope bias was not evident in that no significant differences were found between the slope coefficients (betas) for the two groups ($t = -0.105$, $df = 89$, $P = <0.01$). The coefficients, however, were extremely small (0.048 for second graders and -0.152 for third graders), indicating little relationship between WISC and TONI performance of students.

R^2 values were also extremely small ($R^2 = 0.001$ for second graders, and 0.020 for third graders).

When regressing TONI Quotient scores with those on the WISC Full Scale, it was found that for third graders the slope had a minor negative trend. Thus, children demonstrating a higher level of ability on the TONI appear to actually show relatively lower WISC scores. Differences between the slopes and intercepts of the two groups were not statistically significant, however. Thus homogeneity of regression was evident across the two groups (Reynolds and Brown, 1984) indicating fairness in prediction.

Further, normally-achieving Indian children showed, generally, average to above average performance on the WISC ($X = 123.4$), while their performance on the TONI was at below average levels. Possible explanations which may be considered are:

- (1) Indian children are actually hampered by the non-verbal format.
- (2) The dimension of intelligence measured by the TONI is not placed in as high priority by persons in India as the dimensions measured by the WISC.
- (3) Since the Indian WISC was standardised on a wide cross-section of population, including literate and non-literate subjects, this may have artificially inflated the scores for literate children.

If similar negative regression were to be found for children at higher grade levels, this would imply that educational approaches used in India actually deter children from engaging in the kind of mental activity required to answer TONI items. Research with illiterate children would also be of interest, to determine the relationship of the TONI with schooling effects.

CONCLUSION

"The question of the extent to which basic psychological processes are common to mankind is perhaps the major one being pursued by cross-cultural psychology" (Jahoda, 1980, p. 111). There appears to be no *a priori* criterion by which to distinguish the culture-specific from the universal (Triandis, 1980). Thus, the investigation of psychological processes in each culture of interest warrants research. The present study adds to the literature on the investigation of the above question.

Use of standardised tests developed in the US and Western Europe to determine scholastic aptitude of children in India and other third world countries is gaining popularity with increased interest in special education programme development. However, a careful evaluation of the validity of one instrument reveals that accurate estimation of children's ability cannot be obtained because of bias inherent in the test. If a currently developed test, constructed in adherence to specifications for cross-cultural usage, indicates bias, implications are not promising for other tests in popular usage. Multiple measures, including indigenously developed tests, teacher evaluation, and informal assessment procedures are advised. Exclusive use of abstract-figural tests, and tests not standardised on relevant samples should be avoided.

Correspondence and requests for reprints should be addressed to Dr Rene S. Parmar, Department of Learning and Instruction, 593 Baldy Hall, State University of New York at Buffalo, Buffalo, NY 14260, USA.

REFERENCES

- AMERICAN ASSOCIATION FOR MENTAL DEFICIENCY (1973). Rights of mentally retarded persons. *Ment. Retard.*, 11 (5), 56-58.
- BERK, R. A. (1982). *Handbook of Methods for Detecting Test Bias*. Baltimore: Johns Hopkins University Press.
- BEVIL, U. K. (1982). *Concept Formation in Children: Eco-cultural Context*. New Delhi: Third World Conclave of Social Dynamics.
- BHATIA, C. M. (1955). *Performance Tests of Intelligence under Indian Conditions*. Bombay: Oxford University Press.
- BROWN, L., SHERBENOU, R. J., and JOHNSEN, S. K. (1982). *Test of Non-Verbal Intelligence: A Language-Free Measure of Cognitive Ability*. Austin: Pro-Ed.

- CAMPBELL, D. T., and STANLEY, J. (1966). *Experimental and Quasi-Experimental Designs for Research*. Chicago: Rand-McNally.
- CATTELL, R. B., and CATTELL, A. K. (1959). *Handbook for the Individual or Group Culture-Free Intelligence Test*. Champaign, Ill.: Research Press.
- CHATTERJI, S., and MUKHERJEE, M. (1967). Construction and development of a non-language test of verbal intelligence. *Ind. J. Psychol.*, 46 (3), 235-246.
- CHAWLA, T. R. (1969). An evaluative study of a new culture-free intelligence test (KIT: Exp). *J. psychol. Res.*, 13 (2), 74-76.
- CLARK, P. M. (1985). Review of the Test of Nonverbal Intelligence. In MITCHELL, J. V. (Ed.), *The Ninth Mental Measurements Yearbook: Vol. 2*. Lincoln, Neb.: University of Nebraska Press.
- COOK, T. D., and CAMPBELL, D. T. (1979). *Quasi-experimentation: Design Analysis Issues for Field Settings*. Chicago: Rand-McNally.
- COURT, J. H. (1982). *A Research Bibliography for Raven's Progressive Matrices and Vocabulary Scales*. Bedford Park, South Australia: Flinders University School of Social Sciences.
- CRONBACH, L. J., and DRENTH, P. J. D. (1972). *Mental Tests and Cultural Adaptation*. The Hague: Mouton.
- DAS, J. P., and SINGHA, P. S. (1975). Caste, class, and cognitive competence. *Ind. educ. Rev.*, 10, 1-8.
- DEB, M. (1966). Standardization of a group intelligence test. *Ind. J. Psychol.*, 41 (2), 73-76.
- DESAI, K. G. (1980). Comparative factorial structure of Raven's Standard Progressive Matrices, Cattell's Culture-Fair Scale-3, and Desai-Bhatt Group Tests of Intelligence on samples of various sub-cultures of Gujarat. *J. psychol. Res.*, 24 (1), 8-15.
- DOSAHI, N. L. (1958). Progressive Matrices as a tool for diversification at higher secondary level. *Punjab educ. Q.*, 10, 81-84.
- GHUMAN, A. S. P. (1975). *The Cultural Context of Thinking: A Comparative Study of English and Punjabi Boys*. Windsor: NFER.
- GONZALEZ, E. (1982). Issues in assessment of minorities. In SWANSON, H. L. and WATSON, B. L. (Eds.), *Educational and Psychological Assessment of Exceptional Children*. St. Louis: C. V. Mosby.
- GOODENOUGH, F. L. (1926). *The Measurement of Intelligence by Drawings*. New York: Harcourt-Brace-Jovanovich.
- HUI, C. H., and TRIANDIS, H. C. (1983). Multistrategy approach to cross-cultural research: the case of locus of control. *J. cross-cult. Psychol.*, 14, 65-83.
- HUI, C. H., and TRIANDIS, H. C. (1985). Measurement in cross-cultural psychology. *J. cross-cult. Psychol.*, 16 (2), 131-152.
- HUNDAL, P. S. (1965). A factorial study of verbal and performance test of intelligence. *J. psychol. Res.*, 9 (2), 49-54.
- IRVINE, S. H. (1969). Figural tests of reasoning in Africa: studies in the use of Raven's Matrices across cultures. *Int. Psychol.*, 4, 217-228.
- IRVINE, S. H., and CARROLL, W. K. (1980). Testing and assessment across cultures: issues in methodology and theory. In TRIANDIS, H. C., and BERRY, J. W. (Eds.), *Handbook of Cross-cultural Psychology: Vol. 2. Methodology*. Boston: Allyn and Bacon.
- JACHUCK, K., and MOHANTI, A. K. (1974). Low socio-economic status and progressive retardation in cognitive skills: a test of cumulative-deficiency hypothesis. *Ind. J. ment. Retard.*, 7, 36-45.
- JAHOA, G. (1980). Cross-cultural comparisons. In BORNSTEIN, M. H. (Ed.), *Comparative Methods in Psychology*. Hillsdale, NJ: Lawrence Erlbaum.
- JENSEN, A. R. (1968). Social class and verbal learning. In DEUTSCH, M., KATZ, I., and JENSEN, A. R. (Eds.), *Social Class, Race, and Psychological Development*. New York: Holt, Rinehart and Winston.
- JENSEN, A. R. (1980). *Bias in Mental Testing*. New York: Free Press.
- KAKKAR, S. B. (1975). Linguistic background and intelligence. *J. psychol. Res.*, 19, 132-134.
- MALPASS, R. S., and POORTINGA, Y. H. (1986). Strategies for design and analysis. In LONNER, W. J., and BERRY, J. W. (Eds.), *Field Methods in Cross-cultural Research*. Beverly Hills, CA: Sage.
- MEHROTRA, K. K. (1967). The relationship of the WISC to the Progressive Matrices. *J. psychol. Res.*, 12 (3), 114-118.
- MOHAN, V. (1972). Raven's Standard Progressive Matrices and a verbal test of General Mental Ability. *J. psychol. Res.*, 16 (2), 67-69.
- MOHANTY, B. (1980). Effects of socio-cultural disadvantage on intelligence and short-term memory. *Ind. psychol. Rev.*, 19 (4), 17-24.
- MOHSIN, S. M. (1959). Plea for a scientific aptitude test and a preliminary report on the development of such a test. *Ind. J. Psychol.*, 34, 12-16.
- NAIR, A. S. (1968). *The Kerala University Nonverbal Group Test of Intelligence*. Trivandrum, India: Kerala University Department of Education.
- NAIR, A. S. (1973). Factor analysis of a non-verbal intelligence battery. *J. psychol. Res.*, 17 (3), 121-129.
- NAIR, K. S. (1975). An analytic study of the factor patterns of verbal and non-verbal tests of intelligence. *J. psychol. Res.*, 19 (1), 13-16.
- NANDA, P. C., DAS, J. P., and MISRA, H. K. (1965). Discrimination of geometrical patterns in tribal, rural, and urban children. *J. soc. Psychol.*, 67, 197-200.
- ORTAR, G. (1972). Some principles for adaptation of psychological tests. In CRONBACH, L. J., and DRENTH, P. J. D. (Eds.), *Mental Tests and Cultural Adaptation*. The Hague: Mouton.

- PANDA, K. C., and DAS, J. P. (1970). Acquisition and reversal in four subcultural groups generated by caste and class. *Can. J. Behav. Sci.*, 2, 267-273.
- PATTED, G. M. (1967). *A Study of Spatial Abilities in High School Pupils and its Relation to Success in Architecture Courses*. New Delhi: NCERT.
- PICK, A. D. (1980). Cognition: psychological perspectives. In TRIANDIS, H. C., and BERRY, J. W. (Eds.), *Handbook of Cross-cultural Psychology: Vol. 3*. Boston: Allyn and Bacon.
- POORTINGA, Y. H., and MALPASS, R. S. (1986). Making inferences from cross-cultural data. In LONNER, W. J., and BERRY, J. W. (Eds.), *Field Methods in Cross-cultural Research*. Beverly Hills, CA: Sage.
- RAO, S. (1965). Use of culture free intelligence tests in India. *Ind. J. Psychol.*, 40 (2), 61-66.
- RAO, S. N. (1962). Predicting academic achievement of students in an engineering college. *Psychol. Stud.*, 7, 51-55.
- RAO, S. L., and GUPTA, G. C. (1984). Adaptation of the AH5 test of intelligence to Indian conditions. *J. psychol. Res.*, 28 (2), 59-64.
- RATH, R. (1954). Correlation of Progressive Matrices Test with verbal intelligence and written tests on some other subjects. *Educ. Psychol.*, 1, 4-5.
- RAVEN, J. C. (1960). *Guide to the Standard Progressive Matrices*. London: Lewis.
- RESNICK, L. B., and GLASER, R. (1976). Problem solving and intelligence. In RESNICK, L. B. (Ed.), *The Nature of Intelligence*. Hillsdale, N. J.: Erlbaum.
- REYNOLDS, C. R. (1982). Construct and predictive bias. In BERK, R. A. (Ed.), *Handbook of Methods for Detecting Test Bias*. Baltimore: Johns Hopkins University Press.
- REYNOLDS, C. R., and BROWN, R. T. (1984). *Perspectives on Bias in Mental Testing*. New York: Plenum Press.
- SALVIA, J., and YSSELDYKE, J. E., (1981). *Assessment in Special and Remedial Education*. Boston: Houghton-Mifflin.
- SINGH, A., and HUNDAL, P. S. (1971). Age differences in fluid and crystallized intelligence. *Ind. J. Psychol.*, 46 (1), 85-94.
- SINHA, D. (1973). Priorities and programmes of research in the field of psychology. *J. psychol. Res.*, 17 (1), 22-27.
- SINHA, D. (1978). Story-pictorial EFT: a culturally appropriate test for perceptual disembedding. *Ind. J. Psychol.*, 53, 92-101.
- SINHA, D. (1981). Human assessment in the Indian context. In IRVINE, S. H., and BERRY, J. W. (Eds.), *Human Assessment and Cultural Factors*. New York: Plenum.
- SINHA, M. (1975). Validity of the Progressive Matrices test. *J. psychol. Res.*, 21 (3), 221-226.
- SINHA, M., and CHANDRAKALA, C. (1972). Correlations between Harris' and Burt's Draw-A-Man scales and some other measures of intelligence. *J. psychol. Res.*, 16 (3), 116-120.
- SINHA, R. R. P. (1980). *Tribal and Non-tribal Intelligence*. New Delhi: Classical Publications.
- STERNBERG, R. J. (1980). Sketch of a componential theory of human intelligence. *Behav. Brain Sci.*, 3, 573-584.
- TRIANDIS, H. C. (1980). Introduction. In TRIANDIS, H. C., and LAMBERT, W. W. (Eds.), *Handbook of Cross-cultural Psychology: Vol. 1. Perspectives*. Boston: Allyn and Bacon.
- VERNON, P. E. (1969). *Intelligence and Cultural Environment*. London: Methuen.
- WARHADPANDE, N. R., and SETHI, B. L. (1964). The R-Test of intelligence. *J. psychol. Res.* 8 (1), 19-31.
- WESCHLER, D. (1949). *Wechsler Intelligence Scale for Children*. New York: Psychological Corporation.
- ZAIDI, S. M. H. (1979). Applied cross-cultural psychology: submissions of a cross-cultural psychologist from the third world. In ECKENBERG, L., LONNER, W. J., and POORTINGA, Y. M. (Eds.), *Cross-cultural Contributions to Psychology*. Lisse, The Netherlands: Swets and Zeitlinger.

(Manuscript received 26th September, 1988)

BOOK REVIEWS

ASTINGTON, J. W. *et al.*, (Eds.) (1988). *Developing Theories of Mind*. Cambridge: Cambridge University Press, pp. 447, £10.95 pbk., ISBN 0-415-00595-7.

The ideas, arguments and evidence which pack the 21 chapters of this collection fully justify their publication together. That authors take up points of detail from the other contributions helps to clarify the pattern of similarities and differences in usage of terms and interpretation of data to an extent that is refreshingly novel. The collection is a timely album of photographs of current news about young children's beliefs about minds and their activities. The four sections are entitled: I — Developmental origins of children's knowledge (beliefs?) about the mind, II — Co-ordinating representational states with the world; understanding the relationships among perception, knowledge and reality, III — Further development of a theory of mind: understanding mental states in social interaction and communication, and IV — Further theoretical implications of children's concepts of mind. These give the flavour of the issues, revolving around the young child's passage through the history of the philosophy of mind.

Be warned. The going is hard. The text is dense. The arguments are tight. Just as philosophy books need to be read paragraph by paragraph, with repetition and checking, so do some of the arguments here. With and since Piaget we have begun to face up to the complexities of the details of the intellectual development, and the authors are correctly concerned with precision of both description and explanation.

It is to be regretted if these complexities are reduced to the glib formulae on the dustjacket. To reduce the details to readily assimilable and false generalisations is reminiscent of the degraded and degrading accounts of Piaget's work that made frequent appearances in journal articles as well as student essays. The same over-simplifications should not be encouraged in this brave new field.

Three other regrets can be expressed. The first is the continuing traces of romantic surprise that affects some commentators. We are entitled to see whole children and parts of children as wonderful, but these reactions become dangerous if we begin to see "rapid mastery" or "subtle appreciation" as substitutes for explanations. Related to this is the fact that whether changes occur "quickly" or "slowly", "dramatically" or "gradually", does not explain either how the changes came about or why they did not occur earlier or later than they did. We have to explain changes, but the models implied in a number of chapters are still trapped by Zeno's paradox: they do not address the issue of transitional states.

Zeno does not improve his citation Index count as a result of this publication. Aristotle is the only Ancient Greek to benefit. Kant gains one, Descartes two. Given that philosophers have been writing about some of the conceptual issues in focus in this collection, it is perhaps surprising that more of their contributions are not seen as relevant. It is the systematic conceptual clarifications that are most in danger of being ignored. The historical foundations within psychology itself are not prominent. However, this is not uncommon in vigorous new developments.

This collection registers this vigour and shows that the problems are not intractable. They are a challenge to our intellectual ingenuity. The experiments and results reported show how quickly advances can be made.

W. PETER ROBINSON

DAVIDSON, G. (1988). *Ethnicity and Cognitive Assessment*, Australia: DIT Press, pp. 166, n.p., pbk.

School teachers and academic researchers alike who are au fait with what has come to be associated with the woolly term "multi-cultural education" will welcome yet another contribution from the already impressively extensive range of appropriate publications

from Australia. The editor has brought together in this 166-page collection the thinking, research and personal experience of 16 fellow Australians. The chapters are written primarily for school counsellors operating within the Australian educational system where the use of standardised tests is considerably more widespread than in the U.K. They grow out of a 1985 conference which was concerned with psychological and educational assessment of young "Australians" and children from "ethnic minority groups".

For too long there has been a persistent naiveté about the use of standardised tests of intelligence, creativity, motivation and so on in the increasingly culturally pluralistic nations, despite the warnings of P. E. Vernon in the 1960s. It is therefore encouraging to see here a section devoted to research methodology before proceeding to examine practicalities and outcomes from testing: there is nothing more practical than a good theory, after all.

The recognition of the central role of researchers from "ethnic minority groups" and from "native" Australians (who are they? And why does the word "aborigine" persist?) is the dominant trait in the book's second section where sensible strategies are proposed for trans-cultural teams of practitioners and researchers to work together, assuming that this will help to arrive at more valid assessments of children's scholastic progress in multi-cultural societies. All too often, and possibly throughout the world, children of "minority" groups have been wrongly diagnosed and seen to be defective merely because they were not socialised in the mainstream culture.

But how often does the wheel have to be re-invented? Almost all that had to be said about this was said with authority and much humour by Bernadoni, in his 1964 article in the *Personnel and Guidance Journal*. Intelligent behaviour is unquestionably intelligent only in context, and this is underlined by Baldauf in Chapter 3. The book is refreshingly honest about Australia's record, and sees it in the context of changing international fashion, not least that of "affirmative action".

Sections Two and Three are invaluable for having marshalled so many sound studies about and by "aborigines", but it is all too brief. The accent is on the need to assess cognition within the culture's value system. The mismatch between teachers' language usage and expectations and those of children from "ethnic minorities" is dealt with by Ludwig, and the final section extends the focus on language, touching on second language learning, bilingualism, and language for success in higher education.

BILL TAYLOR

GRANT, D. (1989) *Learning Relations*. London: Routledge, pp. 146, £9.95 pbk., ISBN 0-415-01430-1.

Learning Relations is a book that will be of interest to all those concerned with providing a positive learning environment for young children. It is written in a style that is stimulating and absorbing while dealing with issues that are complex and sometimes controversial. Doreen Grant provides an interesting perspective on how theory can underpin an educational initiative to give it purpose and direction as it develops in relation to varying social phenomena.

Two main issues stand out. One is the recognition and analysis of the problems involved in providing an integrated learning experience for young children in a low income inner city environment. The writings of Freire are the basis for examining the complexity of the relationships that exist among the different parties interested in the education and development of young children. The author examines in some detail the initiatives and strategies that may be used to establish a relationship of trust and acceptance between parents and professionals. Sweeping generalisations are avoided. A sensitive awareness is displayed about the varying life experiences and expectations that parents bring to interactions with teachers and researchers. Rather than seeing this as a problem

Doreen Grant takes it as an opportunity to enable parents to promote their own learning by interacting with other parents in a purposeful non-threatening context. The thesis is that the professionals should act as a resource and facilitators in order that parents can increase the control they have over their own lives, ultimately to the benefit of their children.

The second major dimension to the book is an examination of the importance of the relationship between language and thought. In particular the works of Luria, Bruner and Donaldson are seen as providing appropriate explanations of cognitive development and "self-directed ability". Both the written and spoken word are seen as fundamental in the development of the kind of thinking that will enable parents and children to prosper from their educational and everyday experience. The author states how Freire and Donaldson "showed in very different ways how language and literacy were central to empowering people to participate in decision making and so play a responsible part in society". The first chapter is called "The problem of mismatch" and relates to what are seen as fundamental differences between home and school. Language is seen as the key to resolve the differences.

For readers actively involved in work similar to that discussed in *Learning Relations* the book may be an inspiration and at the same time somewhat daunting. The achievements of Doreen Grant deserve admiration for success under what at times must have seemed almost impossible circumstances. The problems and disappointments that can belay a project of this kind are not hidden from the reader. The fact that the professionals had as much to learn as the parents is made clear. The frustration of dealing with bureaucrats and the lack of resources are well documented.

Despite all these problems the pervading mood of the book is one of optimism. As well as containing constructive insights into the nature of equality and co-operation it communicates the satisfaction when progress is made in this direction. The book offers ways of bringing together the two different worlds that many children experience, particularly for those in low income areas where the difference in attitudes, expectations and assumptions lead to suspicion and distrust of "do-gooders".

JOHN HAWORTH

GURNEY, P. W. (1988) *Self-Esteem in Children with Special Educational Needs*. London: Routledge, pp. 166, £17.95 hbk, ISBN 0-415-00599-X.

There has been renewed interest among educationalists in the self-concept and self-esteem and this has been reflected in a number of recent publications including those by Lawrence (1987) and Robinson and Maine (1988). In his study Peter Gurney has focused on children with special educational needs and his treatment is characterised by empathic understanding, particularly for those youngsters whose difficulties are related to emotional and behavioural disorders.

For its size the book is surprisingly comprehensive in scope. Chapter 1 examines a number of theoretical issues including definitions of basic terms. The four main conceptual perspectives underlying self-esteem studies — psycho-analytical, humanistic, phenomenological and behavioural — are also briefly introduced. In Chapter 2 Gurney moves on to consider developmental aspects and here three stages are identified. It is argued that where children have a developmental lag it is clearly useful to know the level or stage they have reached in order to intervene appropriately.

The focus of Chapter 3 is on assessment and the techniques examined include Q Sort, Repertory Grid Technique, Semantic Differential Free Response Methods and Interviews. Understandably most space is given to the more frequently used self-rating scales. The

problems of reliability and validity in all of these approaches are appropriately underlined. Indeed it is concluded that teachers can probably best assess their children's self-esteem by personally getting to know them as well as possible. A list of behaviours considered to be indicative of low self-esteem, against which the teacher's knowledge can be checked, is provided.

Chapters 4 through to 7 constitute the heart of the book. Here self-esteem is discussed in terms of school and classroom factors and it is maintained that "self-esteem permeates the child's whole life and potentially influences every single learning situation" (p. 51).

Particular attention is given in Chapter 4 to the quality of schooling and academic achievement. Children with learning difficulties have always been seriously disadvantaged in those schools which place most value upon academic achievement. Certainly a current fear among many special educators is that the implementation of the 1988 Education Reform Act will ensure that the number of such academically orientated schools will be greatly increased. The importance of enhancing and maintaining a high level of self-esteem among children with learning difficulties is well established and, therefore, this possible ERA scenario is unwelcome.

Peter Gurney also points to the research findings showing that children with special educational needs tend to have higher self-esteem in separate special schools and classes than they do in mainstream classes. He argues that the quality of educational experiences which the pupils receive is more important than the type of placement and only children who are "psychologically motivated" should be integrated.

The book gives considerable attention to methods for fostering self-esteem among children and Peter Gurney suggests both general strategies and specific techniques which teachers could use. His general suggestions include the development of teacher/child relationships based on mutual acceptance; creating success in schools; counselling; social skill training; classroom contracting and various extra curricular activities. More specific classroom activities include keeping a "diary of good things", developing a record of personal experience/achievement and increasing among children the frequency of positive self-referent verbal statements (PSRVS).

Finally various ways in which self-esteem among teachers may be enhanced is discussed including the provision of relevant INSET and establishing a network of co-operative staff support. More generally, however, high self-esteem among pupils and teachers is seen as being reciprocally related and it is this virtuous circle which "good schools" do so much to encourage.

This is a stimulating and practical book and its value certainly extends beyond the area of special educational needs. It is a pity it is so expensive.

IAN PETRIE

REFERENCES

- LAWRENCE, D. (1987), *Enhancing Self-Esteem in the Classroom*. London: Paul Chapman.
ROBINSON, G., and MAINE, B. (1988) *They can because . . .* A.W.M.C.

HUTT, S. J., TYLER, S., HUTT, C., and CHRISTOPHERSON, H. (1989). *Play, Exploration and Learning*. London: Routledge, pp. 282, £30.00 hbk., ISBN, 0-415-01286-4.

Much of the research from which this monograph is derived was commissioned by the DES in the late 1970s but it is augmented by later studies. The delay in making the major report was partly caused by the untimely death of Corinne Hutt and partly by a two-year stay "on an official desk". Although psychologists may regret the omission of statistical tests, which will be supplied on request, much of the data which are available remain relevant despite this delay.

The report and discussion will also be of interest to pre-school practitioners and educationalists particularly in relation to the current debate about appropriate educational provision for children under statutory school age. The research was carried out in nursery schools and classes, in day nurseries and in playgroups. Their organisation, staff characteristics, particularly play and child/teacher interaction are identified and the implications for practice discussed. There was confirmation of suggestions made by other studies that there is a marked difference between the value placed on fantasy play by most adults in the study and the actual benefits.

Although there were differences between ethos and training of staff in various institutions a "remarkably similar" practice was identified. Much emphasis was put on free play although nursery teachers imposed more organisation. As the authors admit, it is likely that there would be more structured play at the present time. Teachers' and other caregivers' priorities were identified by questionnaire and by teachers ranking given priorities. This method may have been loaded as, for example, "giving mothers time to themselves" is hardly likely to attract while "fostering a child's intellectual development at her own rate" is. Not a remit of this research, it might nevertheless be speculated that the former might be a priority if it were suggested that the enhancement of the quality of life of families and therefore of children was a good reason for pre-school provision. But, except in playgroups, staff, although generally agreeing on their need to know about home background, rarely involved parents in any capacity.

To a large extent, the report confirms much of what is already known or inferred about pre-school provision, which is rehearsed in the House of Commons Select Committee Report on educational provision for the under-5s. What is more interesting as we try to unravel meanings about and definitions of play is the extension of Hutt's (1966) distinction between exploration and play. Later, this distinction (what does this object do, as against what can I do with this object) was relabelled as "epistemic" or "ludic" behaviour because teachers were found to describe almost any activity of the young child at school as play except biological necessities. Stemming from this work and drawing on observation in the study the authors offer a taxonomy which attempts to account for "all the intrinsically motivated self-chosen activities in the pre-school which generically we call play". The taxonomy allows classification of activities into either epistemic or ludic behaviour with games with rules occupying a place between the two. Epistemic behaviour it is suggested is concerned with acquisition of knowledge and information and ludic with self amusement. The authors suggest that free play, or ludic behaviour, prevails in much pre-school education and needs to be balanced by activities focused on epistemic behaviour. Questions arise about the view of learning implicit within the taxonomy and particular classifications which can be debated. For example, practice may be necessary for mastery over materials and is not necessarily wholly ludic. However, the taxonomy may be useful to future research and to reflections on the nature of pre-school experience.

DEIRDRE PETTIT

REFERENCE

HOUSE OF COMMONS (1988). *Education, Science and Arts First Report: Educational Provision for the Under Fives*. London: HMSO.

MANDEL, H. P., and MARCUS, S. I. (1988). *The Psychology of Underachievement: Differential Diagnosis and Differential Treatment*. New York: Wiley, pp. 397, £37.95 ISBN 0-471-84855-7.

This is a book about the relatively unfashionable concept of underachievement. One of the reasons for the decline in activity in this area is the difficulty of a satisfactory definition. For underachievement to be recognised, there first of all has to be an idea of what normal achievement is. In early studies this was based on the subjective judgments of teachers and then discrepancy between standardised measures such as IQ scores and

standardised achievement tests. Despite statistical refinements such as regression analysis, the problems of using IQ scores as a more or less static standard of comparison or prediction have meant that many recent researchers have turned away from this approach.

Not so the authors of this present book. They have been working steadily on their approach which they term the Developmental Theory Model since the 1960s and claim that it provides a consistent framework for teaching, training and research. In their review of the literature in the first three chapters on the definition of underachievement and the variables which influence the achievement process, they do not minimise the difficulties of drawing any consistent conclusions but claim that their model provides one way to explain and integrate many of these inconsistent findings.

In the next three chapters the authors outline their own approach. Again they tread an unfashionable path, making great play of their approach being based on the application of a medical model of diagnostic treatment. This has been used as a favourite whipping boy in the behaviourist literature of the last couple of decades. The present authors stoutly defend such a model and systematically outline an approach to the understanding and treatment of academic underachievement. Their approach is a very clinical one and is aimed at mental health professionals who work individually with clients over a period of time. They make extensive use of the diagnostic interview and provide helpful details of how this may be used to gather information on such things as school performance, family and social relationships, self perceptions and plans for the future from a client. The data are used to provide a differential diagnosis as to which personality style underlies the underachievement. The clinician is helped in this task by the Developmental Theory Model which is described as a diagnostic scheme. This eclectic model draws upon the work of a wide variety of theorists who are seen as having a particular range of convenience. In the stage model presented, Freud is used to understand the Oedipal stage of development at an age 5 to 7 years while Rogers' client centred approach to treatment is seen as most appropriate in late adolescence.

At each stage in this model the authors outline the problems which might arise if the challenges and relationship needs of that stage are not met. They use a standard American psychiatric diagnostic scheme to describe the different personality styles of underachievers which sounds unfamiliar to British readers. They concentrate on four categories which they claim include the majority of underachieving students. These are the over-anxious disorder, the academic problem, the identity disorder and the conduct disorder.

In part four a chapter is devoted to each of these types and in many ways this is the most useful section of the book. They give an account of the general characteristics of each type together with transcripts of interviews and detailed treatment outlines. For this reviewer the chapter on the academic problem underachiever was the most helpful and the extensive examples and guidance provided will be of real practical use.

JOHN THACKER

SUGDEN, D. (Ed.) (1989). *Cognitive Approaches in Special Education*. Basingstoke: Falmer Press, pp. 220, £25, cloth, ISBN 1-8500-419-6; £10.95 pbk., ISBN 1-8500-418-8.

ASHMAN, A., and CONWAY, R. (1989). *Cognitive Strategies for Special Education*. London: Routledge, pp. 288, £25 cloth, ISBN 0-415-00594-9; £10.95 pbk., ISBN 0-415-00595-7.

JONES, N. (Ed.) (1989). *Special Educational Needs Review: Volume 1*. Basingstoke: Falmer Press, pp. 236, £20 cloth, ISBN 1-85000-488-9; £9.95, pbk., ISBN 1-85000-489-7.

David Sugden has carefully compiled readings chosen both for having complementary topics within cognition and special education and for the expertise and lucidity of the writers. The book is one of very few without redundant sections and represents good value as a primer for undergraduate psychology students, trainee educational psychologists or teachers of pupils with special educational needs (SEN). Whilst the readings do not push back new frontiers they do enable practitioners to apply the material which is well capable of ready generalisation.

After two introductory chapters the topics range across SEN and concentrate on "the 20 per cent". Skills generalisation, social context, Instrumental Enrichment, language and communication, working with parents, behaviour, motoric performance and computer-aided learning are among the areas discussed in the context of cognitive approaches.

Ashman and Conway, like Sugden, spend the first two chapters outlining the changes in special education before immersing the reader in cognitive strategies. This book is more technical and prescriptive than Sugden and describes "process-based instruction (PBI)" through examples in the classroom and in wider context such as teacher-training. A realistic attempt is made to synthesise methods validated by research. Assessment, strategy development, inter-tasks transfer and consolidation are subject to review and critique before examples of application are considered.

The latter book will be useful to many experienced psychologists though Sugden's book can provide an excellent primer where this is required.

It is ironic that Sugden and Ashman and Conway give such (welcome) emphasis to process within the curriculum at a time when a product-based National Curriculum is being implemented.

Neville Jones has edited Volume One of a new series of publications from Falmer Press that is, I presume, going to cover most areas of SEN from the perspective of current practice. This first one is in five parts and sets out to cover special needs and integration, learning and parents, behaviour difficulties, teachers and in-service training and "support for special needs". The latter is simply a description of The National Library for the Handicapped Child and The Voluntary Council for Handicapped Children, though both chapters are eminently readable. But how can one purport to deal with "support" with such slim representation?

Most of the writers are the usual well-known ones and this helps to remind the reader one is living on a smallish island. For those who attempt to keep up with current literature this book may be just a distillation of viewpoints previously realised, heard or read. As an introduction to the SEN scene in England and Wales part 1 is admirable. Paired reading is dealt with as part of "Learning and Parents" in a balanced way by Roger Morgan. "Behaviour Difficulties" as a topic is, as usual, covered in a most thorough, sensible and readable way by both Docking and Tattum. What a pity the timing of this publication just missed giving these two writers the chance to comment on the Elton Report.

The fourth part on "Teachers and In-Service Training" does not provide a lot of useful material for practising educational psychologists. The short report by Robson and Sebba does not have space for the fine details of service delivery, particularly for very short school-based courses that the majority of psychologists will contribute to. The chapter on stress in teaching by Dunham will be old-hat to most psychologists, if important to beginning teachers.

Overall this book edited by Neville Jones certainly deserves to be in college and university libraries.

MIKE SMITH

VERMA, G. K., and BAGLEY, C. (1989) *Cross-Cultural Studies of Personality Attitudes and Cognition*. Basingstoke: Macmillan, pp. 217, £33.00 hbk., ISBN 0-333-39539-5.

Coming close on the heels of their *Personality Cognition and Values* and *Self Concept and MultiCultural Education*, this distinguished couple of writers have yet again delighted their readers by bringing together in one volume a collection of substantive and authoritative research reports and commentaries that bring authority and rigour to investigations into psychological aspects of many of the issues that perpetually dominate the thinking of parents, teachers and social workers who are concerned with ensuring equity and understanding in a multi-racial society to the benefit of all individuals and groups of individuals. The "multi-cultural" debate all too often gets clogged up with rhetoric and passion: here the sharpness and lucidity of empirical research and the focus on psychology provide the debate with much new light, adding to its authority by drawing its material from work in a wide range of countries.

In many ways it is a festschrift to P. E. Vernon, who could be thought of as the father of cross-cultural investigation in psychology. It is divided into two parts, the first being concerned with theoretical issues where Paul Kline argues for the special contribution of cross-cultural study to our understanding of personality and the book's editors review the theoretical and practical progress made in the past two decades in cross cultural-research, arguing the need for such investigation as a secure basis for formulating social and educational policy in a truly multi-cultural society.

The second part looks at specific empirical studies undertaken in Canada, the Caribbean, Japan, Europe, the U.S.A., each written in readable style. I found Bagley's chapter on Cognitive Style and Cultural Adaptation especially interesting and although concerned with Canadian society it has many parallels for psychologists and educationists in the U.K. School teachers in multi-racial U.K. will find much of interest in what is said here about an individual's field-dependence being related to factors such as ethnicity, the nature of technology in society, and the amount of formal schooling.

Funded by the Dutch Bernard Van Leer Foundation, Bieschewel offers a perceptive analysis of work in South Africa with very young children. Starting by assuming deficiency, the project provided compensatory education to children of different ethnic groups in South Africa's divided society; Bieschewel claims that disadvantage appears to be more related to socio-economic rather than to ethnic differences, but he admits that those working in the project have no effective basis for cross-ethnic-group comparisons. Bieschewel's last sentence reads that "early learning may (does? should?) make a contribution to the transformation of a plural into a more integrated society". What message for South Africa does this have for those who feel that integration and assimilation are closely linked and that pluralism and diversity are more likely to ensure mutual respect?

BILL TAYLOR

PUBLICATIONS RECEIVED

Books for review have been selected from the following
("n.p." means "no price mentioned")

- ADAMSKI, W., and GROOTINGS, P. (1989). *Youth, Education and Work in Europe*. London: Routledge, pp. 191, £25.00 hbk., ISBN 0-415-00578-7.
- AINSCOW, M., and MUNCY, J. (1989). *Meeting Individual Needs*. London: David Fulton, pp. 149, £10.95 pbk., ISBN 1-85346-033-8.
- BREAKWELL, G. M. (1989). *Facing Physical Violence*. London: Routledge, pp. 97, £5.95 pbk., ISBN 0-901715-95-6.
- CHAMBERLAIN, R. (1989). *Free Children and Democratic Schools*. Basingstoke: Falmer Press, pp. 146, n.p., pbk., ISBN 1-85000-580-X.
- CLARK, M. (1989). *Language through Living for Hearing-Impaired Children*. Sevenoaks: Hodder and Stoughton, pp. 92, £6.95 pbk., ISBN 0-340-50537-0.
- COLLEY, A. M., and BEECH, J. R. (1989). *Acquisition and Performance of Cognitive Skills*. Chichester: John Wiley and Sons, pp. 348, £32.95 hbk., ISBN 0-471-91461-4.
- FISH, J. (1989). *What is Special Education?* Milton Keynes: Open University Press, pp. 112, £8.95 pbk., ISBN 0-335-09535-6.
- FONTANA, D. (1989). *Managing Stress*. London: Routledge, pp. 118, £5.95 pbk., ISBN 0-901715-97-2.
- FRAISSE, P., and SEGUI, J. (1988). *L'Année Psychologique, Fasc. 4*. Paris: Presses Universitaires de France, pp. 159, n.p. pbk., ISBN 2-13-042-036-2.
- GARTON, A., and PRATT, C. (1989). *Learning to be Literate*. Oxford: Basil Blackwell, pp. 249, £8.95 pbk., ISBN 0-631-15835-9.
- GOOCH, S. (1989). *Cities of Dreams*. London: Rider, pp. 278, £16.95 hbk., ISBN 0-7126-1925-9.
- GOODLAD, S., and HIRST, B. (1989). *Peer Tutoring*. London: Kogan Press, pp. 184, £16.95 hbk., ISBN 1-85091-777-9.
- GREER, B., and MULHERN, G. (1989). *New Directions in Mathematics Education*. London: Routledge, pp. 337, £30.00 hbk., ISBN 0-415-00541-8.
- GRUNWELL, P., and JAMES, A. (1988). *The Functional Evaluation of Language Disorders*. London: Croom Helm, pp. 248, £25.00 hbk., ISBN 0-709-95818-8.
- HOWITT, D., et al (1989). *Social Psychology: Conflicts and Continuities*. Milton Keynes: Open University Press, pp. 202, £6.95 pbk., ISBN 0-335-09882-7.
- JOHNSTONE, M. (1989). *Stress in Teaching*. Edinburgh: SCRE, pp. 51, £4.20 pbk., ISBN 0-947833-29-3.
- JONES, N., and SOUTHGATE, T. (1989). *The Management of Special Needs in Ordinary Schools*. London: Routledge, pp. 216, £25.00 hbk., ISBN 0-415-00585-X.
- JOYNSON, R. (1989). *The Burt Affair*. London: Routledge, pp. 347, £25.00 hbk., ISBN 0-415-01039-X.
- MACKAY, G., and DUNN, W. (1989). *Early Communicative Skills*. London: Routledge, pp. 292, £30.00 hbk., ISBN 0-415-03204-0; £12.95 pbk., ISBN 0-415-03205-9.
- MUNRO, A., MANTHEI, B., and SMALL, J. (1989). *Counselling: The Skills of Problem-Solving*. London: Routledge, pp. 161, £5.95 pbk., ISBN 0-415-03956-8.
- POWER, F. C., HIGGINS, A., and KOHLBERG, L. (1989). *Lawrence Kohlberg's Approach to Moral Education*. New York: Columbia University Press, pp. 322, \$40.50 hbk., ISBN 0-231-05976-0.
- PUGH, G., and DE'ARTH, E. (1989). *Working Towards Partnership in the Early Years*. London: National Children's Bureau, pp. 76, n.p. pbk., ISBN 0-902-81742-6.
- RECHLIN, M., and BACHER, F. (1989). *Les Différences Individuelles dans le Développement Cognitif de l'Enfant*. Paris: Presses Universitaires de France, pp. 320, n.p. pbk., ISBN 2-13-042231-4.
- RICHARDSON, A., and RITCHIE, J. (1989). *Letting Go*. Milton Keynes: Open University Press, pp. 93, £6.95 pbk., ISBN 0-335-15840-4.
- SHEPARD, L. A., and SMITH, M. L. (1989). *Flunking Grades: Research and Policies on Retention*. Basingstoke: Falmer Press, pp. 243, £8.95 pbk., ISBN 1-85000-339-4.
- SPENCER, C., BLADES, M., and MORSLEY, K. (1989). *The Child in the Physical Environment*. Chichester: John Wiley and Sons, pp. 302, £32.95 hbk., ISBN 0-471-91235-2.
- SUAREZ-OROZCO, M. M. (1989). *Central American Refugees and U.S. High Schools*. California: Stanford University Press, pp. 182, \$32.50 hbk., ISBN 0-8047-1498-3.
- THARP, R. G., and GALLIMORE, R. (1989). *Rousing Minds to Life*. Cambridge: Cambridge University Press, pp. 317, £25.00 hbk., ISBN 0-521-36234-2.

- WANG, M. C., REYNOLDS, M. C., and WALBERG, H. J. (1989). *Handbook of Special Education: Research and Practice, Vols. 1, 2, & 3*. Oxford: Pergamon Press, pp. 1156, £140·00 hbk.
- WHELDALL, K., and GLYNN, T. (1989). *Effective Classroom Learning*. Oxford: Basil Blackwell, pp. 189, £7·95 pbk., ISBN 0-631-16826-5.
- WOOD, B. L., (1989). *Children of Alcoholism*. New York: New York University Press, pp. 166, \$16·00 pbk., ISBN 0-8147-9222-7.

AUTHOR INDEX TO VOLUME 59, 1989

ARAD, M.: see Bar-Tal, D.	
ARCHER, J.: Gender-Stereotypic Perceptions of Academic Disciplines	306
BABAD, E.: When Less Information is More Informative: Diagnosing Teacher Expectations from Brief Samples of Behaviour.....	281
BAR-TAL, D.: Effects of Information on Student-Teachers' Stereotypic Perception of Pupils	143
BARWICK, J.: Relations between Reading and Musical Abilities	253
BERNIERI, F.: see Babad, E.	
BLATCHFORD, P.: Teacher Expectations in Infant School: Associations with Attainment and Progress, Curriculum Coverage and Classroom Interaction	19
BOREHAM, N. C.: Modelling Medical Decision-Making Under Uncertainty	187
BOWDEN, J.: see Ramsden, P.	
BOYLE, G. J.: Prediction of Academic Achievement Using the School Motivation Analysis Test.....	92
BOYLE, G. J.: Sex Differences in the Prediction of Academic Achievement Using the Children's Motivation Analysis Test.....	245
BROTHERTON, C.: see Ullah, P.	
BUCKHALT, J. A.: The British Ability Scales Speed of Information Processing Subtest: What Does it Measure?.....	100
BURKE, J.: see Blatchford, P.	
CORNISH, I. M.: The Relationship Between Convergence-Divergence and the Free Recall of Discourse.....	258
COWAN, R.: Children's Use of Counting and Guidelines in Judging Relative Number.....	200
DANIELS, H.: see Cowan, R.	
DUERDOTH, P.: see Szaday, C.	
ENTWISTLE, N.: Pupils' Perceptions of School and Teachers.	
I Identifying the Underlying Dimensions	326
II Relationships with Motivation and Approaches to Learning.....	340
ERNEST, P.: Essay Review: Piaget's Epistemology and the Social Practice of Mathematics ..	266
FARQUHAR, C.: see Blatchford, P.	
FRANCIS, L. J.: Monitoring Changing Attitudes Towards Christianity Among Secondary School Pupils Between 1974 and 1986.....	86
FREEDMAN, S.: see Archer, J.	
FURNHAM, A.: Beliefs Concerning Human Nature: Comparing Current Responses with those Gathered in 1945 and 1956.....	1
GLEESON, D.: Hyperactivity in a Group of Children Referred to a Scottish Child Guidance Service: A Significant Problem.....	262
GOLDCHSTEIN, M.: see Van Overwalle, F.	
GRANLEESE, J.: Teachers' and Boys' and Girls' Perceptions of Competence in the Primary School: the Importance of Physical Competence	31
HALL, E. J.: see Boyle, G. J. (first entry)	
HARPER, G.: Interpretation of Factor Analyses from the Approaches to Studying Inventory ..	66
HONESS, T. M.: School Leavers' Changing Identifications with Different Social Categories and their "Internalisation" of the Category "School Leaver"	59
HOUGHTON, S.: see Wheldall, K.	
HOWE, M.: Separate Skills or General Intelligence: The Autonomy of Human Abilities	351
HUNTER, J.: Cross-Validation of Short Forms of the WISC-R in Two British Samples.....	366
JAEGER, M.: see Norwich, B.	
JENSEN, A. R.: see Buckhalt, J. A.	
KEK, L. Y.: see Townsend, M. A. R.	

- KEMBER, D.: see Harper, G.
- KOZÉKI, B.: see Entwistle, N.
- LANSDOWN, R.: see Hunter, J.
- LEACH, C.: see Lister, C. (first entry)
- LEACH, C.: see Lister C. (second entry)
- LISTER, C.: Similar-Sequence and Similar-Structure in Retarded and Non-Retarded Children's Development..... 8
- LISTER, C.: The Development of Conservation Concepts in Children with Visual Impairments..... 21
- LONKA, K.: Why Does the Length of an Essay-Type Answer Contribute to Examination Marks?..... 220
- LOW, R.: Detection of Missing and Irrelevant Information Within Algebraic Story Problems..... 296
- LYNN, R.: A Nutrition Theory of the Secular Increases in Intelligence: Positive Correlations between Height, Head Size and IQ 372
- McGRAW, D.: see Lister, C. (first entry)
- MARSH, H. W.: The Transition from Single-Sex to Co-educational High Schools: Teacher Perceptions, Academic Achievement and Self-Concept 155
- MARTIN, E.: see Ramsden, P.
- MAY, D. S.: see Worrall, N.
- MERRETT, F.: see Wheldall, K.
- MIKKONEN, V.: see Lonka, K.
- MOIR, J.: Structured Role Orientation and Academic Self-Concept in the Secondary School..... 241
- MYERS, M. R.: see Marsh, H. W.
- NORWICH, B.: The Predictive Relationship Between Beliefs, Attitudes, Intentions and Secondary School Mathematics Learning: A Theory of Reasoned Action Approach 314
- OVER, R.: see Low, R.
- OWENS, L.: see Marsh, H. W.
- PARKER, D.: see Gleeson, D.
- PARMAR, R. S.: Cross-Cultural Transfer of Non-Verbal Intelligence Tests: An (In) Validation Study..... 378
- PHILLIPS, C. J.: Children's Learning Skills: A Cautionary Note on Ethnic Differences..... 108
- PICKERING, D.: see Szaday, C.
- PLEWIS, I.: see Blatchford, P.
- RAMSDEN, P.: School Environment and Sixth Form Pupils' Approaches to Learning 129
- RAVIV, A.: see Bar-Tal, D.
- ROSENTHAL, R.: see Babad, E.
- SEGEBARTH, K.: see Van Overwalle, F.
- SIMPSON, L.: see Lister, C. (first entry)
- SMITH, I. D.: see Marsh, H. W.
- START, K. B.: see Boyle, G. J. (first entry)
- START, K. B.: see Boyle, G. J. (second entry)
- SZADAY, C.: One in Eleven: Pupils Requiring Special Educational Assistance in Catholic Schools in the Australian State of Victoria..... 361
- TAIT, H.: see Entwistle, N. J.
- TIZARD, B.: see Blatchford, P.
- TODMAN, J.: see Moir, J.

- TOWNSEND, M. A. R.: The Effect of Mood on the Reliability of Essay Assessment.....232
- TREW, K.: see Granleese, J.
- TUCK, B. F.: see Townsend, M. A. R.
- TURNER, I.: see Granleese, J.
- ULLAH, P.: Sex, Social Class and Ethnic Differences in the Expectations of Unemployment
and Psychological Well-Being of Secondary School Pupils in England.....49
- URBANOWICZ, M. A.: see Hunter J.
- VALENTINE, E.: see Barwick, J.
- VAN OVERWALLE, F.: Improving Performance of Freshmen Through Attributional
Testimonies from Fellow Students.....75
- WALSH, M.: see Lister, C. (second entry)
- WEST, R.: see Barwick, J.
- WHELDALL, K.: Natural Rates of Teacher Approval and Disapproval in British Secondary
School Classrooms.....38
- WILDING, J.: see Barwick, J.
- WORRALL, N.: Towards a Person-in-Situation Model of Teacher Stress 174
- YULE, W.: see Hunter, J.

EDITORIAL ACKNOWLEDGMENT

The Editor and Assistant Editors of *The British Journal of Educational Psychology* gratefully acknowledge the assistance of the following, who have refereed papers submitted between 1st April 1988 and 31st March 1989:

C. Antaki	P. Gurney	F. Potter
N. Austin	P. Hannon	Linda Pring
N. Bennett	R. Harrison	Rea Reason
G. Brown	R. J. Hartley	I. St. James Roberts
P. Bryant	M. J. Howe	C. Rogers
Anne Cockburn	M. Hughes	A. Salisbury
C. Cooper	C. Hulme	A. Seddon
T. Cox	P. N. Johnson-Laird	J. Stevenson
Alyson Davis	P. Kline	J. Thomas
A. Dobbins	C. Kyriacou	P. Tomlinson
Rosalind Driver	P. M. Levy	J. Versey
N. J. Entwistle	Flora J. MacLeod	K. Wheldall
P. Ernest	J. Ogborn	N. Worrall
N. Freeman	Deirdre Pettitt	E. C. Wragg
W. Gillham	H. Pont	M. Youngman

370.1505
B777

NOTES TO CONTRIBUTORS

Manuscripts are accepted on the understanding that they report unpublished work that is not under consideration for publication elsewhere, and that if accepted for the *Journal*, the work will not be published again in the same form, in any language, without the consent of the Editor.

The usual range of length of papers is from 2,500 to 4,000 words. Papers should be in a clear style and free from jargon. Statistical matter (formulae, tables, discussion of techniques, etc.) should be kept to a minimum, so that the *Journal* will be of interest to statistically untrained readers. Such concepts as mean, percentiles, standard deviations, correlation, chi-squared, standard error, critical ratio, significance and reliability, regression and factor analysis, can be employed in the text and need not be explained.

Short research notes will also be published, reporting work of two kinds: (i) results of sufficient moment to merit publication in advance of a more comprehensive paper; (ii) work which substantially confirms or extends existing knowledge but which does not justify an extensive paper.

Contributors can help to avoid delays in publication by observing certain guiding principles:

1—Papers should be written with the utmost conciseness consistent with clarity. If the editor or a referee takes the wrong meaning from or fails to understand a passage, the fault is the author's. Unless the editor can grasp the meaning of a sentence unequivocally, it cannot be assumed that other readers can.

2—A paper should be written only when a piece of work is rounded off. A report of a pilot study, or a series of papers on the same subject written as the results come to hand, is unlikely to be accepted for publication. A comprehensive paper, particularly if it includes reports of several experiments or analyses, or replications on further samples, which give consistent results, will be given favourable consideration.

3—Experimental reports should generally follow the standard pattern of presentation, e.g.:

- (a) A short summary at the start.
- (b) A statement of aims or objectives and a review of relevant previous work as an introduction to the topic — conciseness and strict selection of relevant aspects are particularly important here.
- (c) A clear statement of the sample, the procedure, the tests, the experimental material, in sufficient detail to permit a reader to replicate the study if he so wishes.
- (d) A factual statement of results, without speculative comment or discussion.
- (e) Either a brief list of general conclusions, not repeating results already listed, or a discussion, which should be strictly limited to discussing the results (if this is necessary) and which should not involve mere recapitulation or speculation beyond the scope of the inquiry.
- (f) Acknowledgments.
- (g) References.

4—The summary (which is often copied by abstracting agencies) should be intelligible without reference to the paper. It should give an indication of the aim, scope and results of the inquiry, preferably in about three sentences. It should mention the size and nature of the sample, and should not include vague phrases such as "the results were analysed . . ." or "The implications of the findings are discussed . . .". Even for a full-length paper, the summary should not exceed 200 words.

5—Figures and tables should be selected to illustrate or summarise points made in the text. The same data should not be reported both in table and figure; tables are generally preferable because they are much more economical. Tables should contain selected data: only the most important quantitative results can be reported in full.

6—Authors should consult a current issue of the *Journal* in order to follow the conventions applied in the *Journal* — the use of headings, numbering of tables, the absence of footnotes, the impersonal style of reporting, omission of dots in standard abbreviations, reference to subjects as children and not as Ss, the lay-out of references and so on. The British Psychological Society pamphlet, *Suggestions to Contributors*, should be consulted. References should be cited by author and date in the text, and arranged in alphabetical order in the list of references at the end. All measurements should be stated in SI units. The *Journal* is committed to the use of non-sexist language.

7—A manuscript should be headed with the title of the paper, the names of the authors and, in brackets, a short address identifying the institution where the work was done. Titles should be short. If the title exceeds 45 letters and spaces in length, a shortened version, suitable as a running title, should be added.

8—The manuscript should be typed in double-spacing with wide margins on one side of the paper which is suitable for ink corrections and printer's instructions — not duplicated or on flimsy paper. Drawings should be in Indian ink on cartridge paper. Manuscripts which require retyping before they can be handled by the printer have to be sent back to the authors. The *Journal* has no paid staff.

9—Authors, if they specifically request, may have their papers reviewed blind.

Authors receive 50 copies of their papers free (delivered about six weeks after publication). Extra copies may be had at cost price if the order is given when the proof is returned.

The *Journal* will adopt the International System of Units (SI) based on the units: metre, kilogramme, second, ampere, Kelvin and candela. Authors who refer to physical measurements in their papers should now use SI units except where it would be absurd to do so. Further information about SI units, including conversion tables, is contained in the revised *Suggestions to Contributors* pamphlet issued by the British Psychological Society and obtainable, price £2.00 post free, from The British Psychological Society, St Andrew's House, 48 Princess Road East, Leicester LE1 7DR.

Reading and Writing: An Interdisciplinary Journal

Editor:

R. Malatesha Joshi, *Dept. of Reading, School of Education and Human Development, Fayetteville State University, NC, USA*

Reading and writing skills are fundamental to literacy. Consequently, the processes involved in reading and writing and the failure to acquire these skills, as well as the loss of once well developed reading and writing abilities have been targets of intense research inquiry. This has involved professionals from diverse disciplines such as neuropsychology, cognitive psychology, psycholinguistics, and education. The research findings that have emanated from these disciplines are written up in a lingua specific to the particular discipline and are published in specialized journals. This generally leaves the expert in one area almost unaware of what is taking place in another area. Even though both areas are concerned with reading and writing. The purpose of this journal is to cut across these artificial boundaries of several disciplines. The journal will focus on the interaction among various fields such as linguistics, information processing, neuropsychology, cognitive psychology, speech and hearing sciences, and education.

The journal will publish research articles, critical reviews, theoretical papers, case studies, and book reviews. Special emphasis will be given to articles that address models, diagnosis and remediation of reading, writing, and spelling; orthography and its relation to reading and writing; computer literacy; cross cultural studies; and developmental and acquired disorders of reading and writing.

Subscription Information **ISSN 0922-4777**
1989, Volume 1 (4 issues)
Institutional rate: Dfl. 236.00/US\$115.50 *incl. p&h*
Private rate: Dfl. 100.00/US\$ 49.00 *incl. p&h*

Private subscriptions should be sent direct to the publisher
Back volume information is available upon request

P.O. Box 322, 3300 AH Dordrecht, The Netherlands
P.O. Box 358, Accord Station, Hingham, MA 02018-0358, U.S.A.

**New
Publication**

**KLUWER
ACADEMIC
PUBLISHERS**



British Journal of Developmental Psychology

Special issue on 'The child's theory of mind'

Submissions are invited for a special issue of the *British Journal of Developmental Psychology* on 'The child's theory of mind'. Empirical and theoretical contributions concerning the development of the child's knowledge of her own mind and the minds of other people are requested. The topic may be approached from any developmental or comparative perspective and may include aspects of emotional development or the developmental psychopathology of metacognition. Papers not chosen for the special issue will be considered for the journal in the usual way. The issue will be published as part of volume 9 (1991).

Papers should be submitted to the Editor, Professor G. E. Butterworth, Journals Office, 13A Church Lane, London N2 8DX, UK *before 28 February 1990*.

The *British Journal of Developmental Psychology* is published four times a year. Subscription price of volume 7 (1989) is £55.00 (US\$103).

A special individual price of US\$54.00 is available to members of APA/SRCD/ISSBD.

The British Psychological Society

The Distribution Centre, Blackhorse Road, Letchworth,
Herts SG6 1HN, UK.

British Journal of Psychology

Edited by Professor D. E. Blackman
(University of Wales College of Cardiff)

Selected contents of volume 80 (1989)

- When is an object not an object? The effect of 'meaning' upon the copying of line drawings — M. Lee
- Verbal operant conditioning as a function of extraversion and reinforcement — S. Gupta & A. P. Shukla
- The memory structure of autistic idiot-savant mnemonists — N. O'Connor & B. Hermelin
- Accurate non-word spelling despite congenital inability to speak — D. V. M. Bishop & J. Robson
- A generation effect in memory without awareness — J. M. Gardiner
- Children's performance in the selection task — V. Girotto, M. Gilly, A. Blaye & P. Light
- A longitudinal study of gender and cultural variations in school leavers' perceptions and experiences of employment and their self-attributions and modes of organizing time — T. M. Honess
- The effect of spaced practice on explicit and implicit memory — P. Perruchet
- Does hand proficiency determine hand preference? — D. V. M. Bishop
- Repetition priming of face recognition in a serial choice reaction-time task — T. Roberts & V. Bruce
- Observer agreement calculations — A. Harrop, C. Foulkes & M. Daniels

Subscription price for volume 80 (4 parts) is £89.00 (US\$169)

The British Psychological Society

The Distribution Centre, Blackhorse Road, Letchworth, Herts.
SG6 1HN, UK

Journals of The British Psychological Society

British Journal of Psychology

Edited by A.J. Chapman, University of Leeds

- ✓ The Journal publishes reports of empirical studies likely to further our understanding of psychology;
- ✓ critical reviews of the literature; and theoretical contributions. Readership is world-wide and
- ✓ includes academic and research workers in experimental and cognitive psychology.
- ✓ Volume 81 (1990) (4 parts) £97.00 (US\$181.00)

British Journal of Medical Psychology

Edited by John Birchall, Institute of Psychiatry

- ✓ The British Journal of Medical Psychology publishes original contributions to knowledge in the area of those aspects of psychology applicable to medicine and related clinical disciplines, including the psychotherapies and counselling.
- ✓ Published articles include empirical studies, theoretical papers and clinical reports which suggest increased understanding of the psychological aspects of medicine.
- ✓ The readers are medical psychologists, mainly clinicians, concerned with applying psychology to medical problems, from the disciplines of psychology, medicine and professions allied to medicine.
- ✓ Volume 63 (1990) (4 parts) £76.00 (US\$141.00)

British Journal of Social Psychology

Edited by Miles Hewstone, University of Bristol

- ✓ The British Journal of Social Psychology publishes articles which contribute to the basic methodological and theoretical issues confronting the discipline. In addition, it encourages: theoretical and review papers; papers describing applied social psychology in a variety of settings; correspondence on controversial issues; guest-edited issues on research in growth areas of social psychology.
- ✓ Volume 29 (1990) 4 parts £65.00 (US\$121.00)

British Journal of Developmental Psychology

Edited by George Butterworth, Professor of Psychology, University of Stirling

- ✓ This new journal, launched in 1983, publishes authoritative scientific articles describing all aspects of development from infancy to old age, including educational implications and the problems of abnormal development.
- ✓ Volume 8 (1990) (4 parts) £61.00 (US\$112.00)

The British Psychological Society

- ✓ The Distribution Centre, Blackhorse Road, Letchworth, Herts SG6 1HN, UK.

British Journal of Mathematical and Statistical Psychology

Edited by A.D. Lovie, University of Liverpool

- ✓ The Journal publishes articles relating to any areas of psychology which have a greater mathematical or statistical or other formal aspect to their argument than is usually acceptable to other journals. Articles with a clear reference to substantive psychological issues are preferred, e.g. new models for psychological processes, new approaches to existing data, critiques of existing models and improved algorithms for estimating the parameters of a model.
- ✓ Volume 43 (1990) (2 parts) £64.00 (US\$119.00)

Journal of Occupational Psychology

Edited by Michael West, University of Sheffield

- ✓ An international journal of industrial and organizational psychology which describes and interprets new research into people at work. Areas covered include personnel and vocational psychology; occupational social psychology; human factors and engineering and computer psychology; organizational psychology; behavioural aspects of industrial relations. Innovative or interdisciplinary approaches to problems are particularly prominent.
- ✓ Volume 63 (1990) (4 parts) £64.00 (US\$119.00)

British Journal of Clinical Psychology

Edited by Ray Hodgson, Department of Clinical Psychology, Whitchurch Hospital, Cardiff

- ✓ The British Journal of Clinical Psychology publishes new findings, theoretical, methodological and review papers bearing on the whole field of clinical psychology and includes: descriptive and aetiological studies of psychopathology; applications of psychology to medicine and health care; studies of the assessment and treatment of psychological disorders; social and organizational aspects of psychological disorders and ill-health. Contributions combine scientific excellence with relevance to the concerns of practising clinicians, bridging the gap between academic and professional interests.
- ✓ Volume 29 (1990) (4 parts) £65.00 (US\$121.00)

All seven journals ordered at the same time for one address £433 (US\$805)



New titles of interest from Springer

B. J. Zimmerman, City University of New York;
D. H. Schunk, University of North Carolina (Eds.)

Self-Regulated Learning and Academic Achievement *Theory, Research, and Practice*

1989. Approx. 255 pp. 5 figs. (Springer Series in
Cognitive Development). Hardcover DM 98,-.
ISBN 3-540-96934-9

In contrast to previous ability or environmental
formulations that address the **why** of achieve-
ment, self-regulation models focus on **how**
students activate, alter, and sustain their learning
practices using a variety of self-related processes.
This book brings together researchers represent-
ing different theoretical perspectives.

M. Crawford, West Chester University;
M. Gentry, Hamilton College (Eds.)

Gender and Thought *Psychological Perspectives*

1989. Approx. 155 pp. 5 figs. Hardcover DM 76,-.
ISBN 3-540-96891-1

A distinguished group of authors shows how
various forces in language, family practices, and
education reinforce differentiation among the sex
roles. This volume attempts to bridge this gap
between difference and equality by revitalizing
and reinterpreting the concept of gender differ-
ences.

C. B. McCormick, G. E. Miller, University of
South Carolina; M. Pressley, University of
Western Ontario (Eds.)

Cognitive Strategy Research *From Basic Issues to Educational Applications*

1989. X, 340 pp. 23 figs. Hardcover DM 98,-.
ISBN 3-540-96869-5

Highly regarded experts review the state of the
art in cognitive strategy research with an
emphasis on the transition from laboratory to
educational contexts.

D. B. McLeod, Washington State University;
V. M. Adams, University of Georgia (Eds.)

Affect and Mathematical Problem Solving *A New Perspective*

1989. X, 268 pp. 16 figs. Hardcover DM 74,-.
ISBN 3-540-96924-1

Research on cognitive aspects of mathematical
problem solving has made great progress in
recent years, but the relationship of affective
factors to problem-solving performance has been
a neglected research area.
The purpose of this book is to show how the
theories and methods of cognitive science can be
extended to include the role of affect in mathe-
matical problem solving.

R. L. Taylor, L. Sternberg, Florida Atlantic
University

Exceptional Children *Integrating Research and Teaching*

1989. X, 419 pp. 23 figs. Hardcover DM 118,-.
ISBN 3-540-96913-6

This unique text is a comprehensive,
graduate-level introduction to special education.
Each chapter includes sections on the adaptation
of teaching procedures, teaching content and
materials, the instructional environment, and
equipment.

Prices are subject to change without notice.

Springer-Verlag Berlin
Heidelberg New York London Paris
Tokyo Hong Kong

Heidelberger Platz 3, D-1000 Berlin 33 · 175 Fifth Ave., New York,
NY 10010, USA · 8 Alexandra Rd., London SW19 7JZ, England ·
26, rue des Carmes, F-75005 Paris · 37-3, Hongo 3-chome,
Bunkyo-ku, Tokyo 113, Japan · Citicorp Centre, Room 1603, 18
Whitfield Road, Causeway Bay, Hong Kong

lm.3250/SF

Springer



PUBLICATIONS OF THE BRITISH PSYCHOLOGICAL SOCIETY

British Journal of Psychology
British Journal of Medical Psychology
Journal of Occupational Psychology
British Journal of Social Psychology
British Journal of Clinical Psychology
British Journal of Developmental Psychology
British Journal of Mathematical and Statistical Psychology
The British Journal of Educational Psychology

The *British Journal of Psychology*, the *British Journal of Medical Psychology*, *Journal of Occupational Psychology*, the *British Journal of Developmental Psychology*, the *British Journal of Social Psychology* and the *British Journal of Clinical Psychology* appear in quarterly parts. The 1989 subscription prices, payable in advance, are £89.00 (US\$169.00) for the *British Journal of Psychology*; £59.00 (US\$113.00) for the *British Journal of Social Psychology*; £59.00 (US\$113.00) for the *British Journal of Clinical Psychology*; £55.00 (US\$103.00) for the *British Journal of Developmental Psychology*; £70.00 (US\$134.00) for the *British Journal of Medical Psychology*; and £59.00 (US\$113.00) for the *Journal of Occupational Psychology*.

The *British Journal of Mathematical and Statistical Psychology* is issued twice a year, in May and November. The subscription rate, payable in advance, is £59.00 per volume post free (US\$113.00 overseas).

Subscriptions to any of these journals may be sent to a bookseller or subscription agent, or to The British Psychological Society, The Distribution Centre, Blackhorse Road, Letchworth, Herts., SG6 1HN, UK. A special price of £403.00 (US\$766.00) for all seven journals going to the same address is offered for 1989.

The *British Journal of Educational Psychology* is issued three times a year, in February, June and November. The prepaid subscription is £25.00 a year post free, or £8.50 each number. The prepaid subscription price in North America is \$50.00 US. Journals are sent accelerated surface post to the USA. Subscriptions should be sent to the Scottish Academic Press, 139 Leith Walk, Edinburgh, EH6 8NS.

MANUSCRIPTS

For the *British Journal of Psychology* should be sent to Professor D. E. Blackman, Journals Office, 13A Church Lane, London, W2 8DX.

For the *British Journal of Medical Psychology* should be sent to Dr Sidney Crown, Department of Psychiatry, The London Hospital, Whitechapel, E1 1BB.

For the *Journal of Occupational Psychology* should be sent to Dr D. Guest, Journals Office, 13A Church Lane, London, W2 8DX.

For the *British Journal of Social Psychology* should be sent to Dr A. S. R. Manstead, Journals Office, 13A Church Lane, London, W2 8DX.

For the *British Journal of Clinical Psychology* should be sent to Professor R. Hodgson, Journals Office, 13A Church Lane, London, W2 8DX.

For the *British Journal of Mathematical and Statistical Psychology* should be sent to Dr A. D. Lovie, Journals Office, 13A Church Lane, London, W2 8DX.

For the *British Journal of Developmental Psychology* should be sent to Professor George Butterworth, Journals Office, 13A Church Lane, London, W2 8DX.

For *The British Journal of Educational Psychology* should be sent in duplicate to Professor Charles Desforges, University of Exeter, St Luke's, Heavitree Road, Exeter, EX1 2LU.

Contributors are advised to consult Suggestions to Contributors issued by the Standing Committee on Publications of the British Psychological Society and obtainable, price £2.00 post free, from The British Psychological Society, St Andrew's House, 48 Princes Road East, Leicester, LE1 7DR.

All manuscripts should be submitted in duplicate.

OFFPRINTS

Contributors on application receive fifty copies of their papers free. Additional copies may be obtained in units of fifty copies.

REPRINTS

The *Journal* is available in reprint from Swets and Zeitlinger B.V., P.O. Box 810, 2160 SZ Lisse, The Netherlands; Vols. 21-29, 31-47.

SUBSCRIPTIONS TO THE BRITISH PSYCHOLOGICAL SOCIETY

The subscription for Fellows, Associates and Graduate Members is £40.00.

They are entitled to purchase any or all of the Society's Journals at the special rates of £8.00 per Journal per annum.

Information on membership and activities may be obtained from the Secretary, The British Psychological Society, St Andrew's House, 48 Princess Road East, Leicester, LE1 7DR.

BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY,
VOL. 59, PART 3, 1989

CONTENTS

ARTICLES

- | | | |
|---|--|------------|
| BABAD, E., BERNIERI, F.,
and ROSENTHAL, R. | When Less Information is More Informative: Diagnosing
Teacher Expectations from Brief Samples of
Behaviour..... | 281 |
| Low, R., and OVER, R. | Detection of Missing and Irrelevant Information Within
Algebraic Story Problems | 296 |
| ARCHER, J., and FREEDMAN,
S. | Gender-Stereotypic Perceptions of Academic
Disciplines..... | 306 |
| NORWICH, B., and JAEGER,
M. | The Predictive Relationship Between Beliefs, Attitudes,
Intentions and Secondary School Mathematics Learning:
A Theory of Reasoned Action Approach | 314 |
| ENTWISTLE, N. J., KOZEKI,
B., and TAIT, H. | Pupils' Perceptions of School and Teachers:
I Identifying the Underlying Dimensions.....
II Relationships with Motivation and Approaches
to Learning..... | 326
340 |
| HOWE, M. J. A. | Separate Skills or General Intelligence: The Autonomy
of Human Abilities | 351 |

RESEARCH NOTES:

- | | | |
|--|---|-----|
| SZADAY, C., PICKERING, D.,
and DUERDOTH, P. | One in Eleven: Pupils Requiring Special Educational
Assistance in Catholic Schools in the Australian State
of Victoria..... | 361 |
| HUNTER, J., YULE, W.,
URBANOWICZ, and
LANSDOWN, R. | Cross-Validation of Short Forms of the WISC-R in two
British samples..... | 366 |
| LYNN, R. | A Nutrition Theory of the Secular Increases in Intelligence:
Positive Correlations between Height, Head Size and
IQ | 372 |
| PARMAR, R. S. | Cross-Cultural Transfer of Non-Verbal Intelligence Tests:
An (In) Validation Study | 378 |

BOOK REVIEWS.....	389
PUBLICATIONS RECEIVED	397
AUTHOR INDEX FOR VOLUME 59.....	399
EDITORIAL ACKNOWLEDGMENT	402
NOTES TO CONTRIBUTORS	403